



UNIVERSIDADE ESTADUAL  
DE CAMPINAS

INSTITUTO DE MATEMÁTICA ESTATÍSTICA  
E COMPUTAÇÃO CIENTÍFICA

**Viviana Analia Ramirez**

# **Reformulação de um problema de programação não linear com restrições multiobjetivo**

CAMPINAS

2015

VIVIANA ANALIA RAMIREZ

**Reformulação de um problema de programação não linear com restrições  
multiobjetivo**

Tese apresentada à Instituto de Matemática  
Estatística e Computação Científica da  
Universidade Estadual de Campinas  
como parte dos requisitos exigidos para  
a obtenção do título de Doutora em  
Matemática Aplicada.

ORIENTADOR: Roberto Andreani

COORIENTADORA: Sandra Augusta Santos

O ARQUIVO DIGITAL CORRESPONDE À VERSÃO  
FINAL DA TESE DEFENDIDA PELA ALUNA VIVIANA  
ANALIA RAMIREZ, E ORIENTADA PELO PROF.  
DR. ROBERTO ANDREANI.

A handwritten signature in black ink, appearing to be 'R. Andreani', is written over a horizontal line.

CAMPINAS

2015

**Agência(s) de fomento e nº(s) de processo(s):** CNPq

Ficha catalográfica  
Universidade Estadual de Campinas  
Biblioteca do Instituto de Matemática, Estatística e Computação Científica  
Maria Fabiana Bezerra Muller - CRB 8/6162

R145r Ramirez, Viviana Analia, 1976-  
Reformulação de um problema de programação não linear com restrições multiobjetivo / Viviana Analia Ramirez. – Campinas, SP : [s.n.], 2015.

Orientador: Roberto Andreani.

Coorientador: Sandra Augusta Santos.

Tese (doutorado) – Universidade Estadual de Campinas, Instituto de Matemática, Estatística e Computação Científica.

1. Programação não-linear. 2. Programação multiobjetivo. 3. Otimização em dois níveis (Matemática). I. Andreani, Roberto, 1961-. II. Santos, Sandra Augusta, 1964-. III. Universidade Estadual de Campinas. Instituto de Matemática, Estatística e Computação Científica. IV. Título.

Informações para Biblioteca Digital

**Título em outro idioma:** Reformulation of a nonlinear programming problem with multiobjective constraints

**Palavras-chave em inglês:**

Nonlinear programming

Multiobjective programming

Bilevel optimization (Mathematics)

**Área de concentração:** Matemática Aplicada

**Titulação:** Doutora em Matemática Aplicada

**Banca examinadora:**

Roberto Andreani [Orientador]

José Mario Martínez Pérez

Maria Aparecida Diniz Ehrhardt

Graciela Noemi Sottosanto

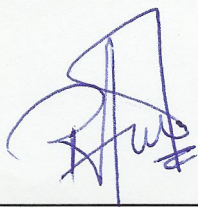
Geraldo Nunes Silva

**Data de defesa:** 30-09-2015

**Programa de Pós-Graduação:** Matemática Aplicada

**Tese de Doutorado defendida em 30 de setembro de 2015 e aprovada**

**Pela Banca Examinadora composta pelos Profs. Drs.**



---

**Prof(a). Dr(a). ROBERTO ANDREANI**



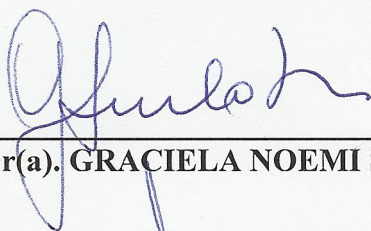
---

**Prof(a). Dr(a). JOSÉ MARIO MARTÍNEZ PÉREZ**



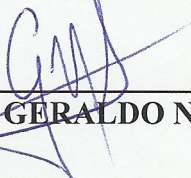
---

**Prof(a). Dr(a). MARIA APARECIDA DINIZ EHRHARDT**



---

**Prof(a). Dr(a). GRACIELA NOEMI SOTTOSANTO**



---

**Prof(a). Dr(a). GERALDO NUNES SILVA**

*“Não temas, porque eu sou contigo;  
não te assombres, porque eu sou o teu Deus;  
eu te fortaleço, e te ajudo,  
e te sustento com a minha destra fiel”.*

*Isaías 41:10*

# Dedicatória

Este trabalho é dedicado a minha amada família.

A meu esposo, companheiro e amigo Emilio.

A minhas filhas, as princesinhas que trouxeram tanta alegria

a nosso lar, Candela e Antonella.

A meus pais Omar e Nidia e a meu irmão Gustavo.

# Agradecimentos

Primeiramente agradeço a Deus, por sua fidelidade, força e ajuda em cada um destes anos, sem Ele não poderia ter concluído este trabalho.

Agradeço a meu amado esposo Emilio, por seu apoio, paciência, amor e a sua ajuda cada vez que precisei.

Agradeço às princesas de nosso lar, as minhas amadas filhas Candela e Antonella. À Candela agradeço seu amor e sorrisos que tanta força me deram em momentos difíceis e à Antonella, a felicidade que trouxe com a sua chegada a nosso lar.

A meus pais Omar e Nidia e a meu irmão Gustavo, que por mais longe que estejam, tenho todo o seu apoio e afeto.

A Nino e Sandra, pelo apoio e confiança que tiveram em mim para realizar este trabalho. Por suas orientações e cada uma das suas sugestões que ajudaram para concluir este trabalho.

Agradeço ao professor José Mario Martínez, por suas sugestões que enriqueceram muito este trabalho.

Agradeço ao Luis Felipe, pelas suas sugestões para este trabalho.

A todos os professores que nestes anos de uma ou outra forma me brindaram parte de seu conhecimento.

A meus amigos Cleyson, Márcia, Elisa e Débora, por terem me dado hospedagem e apoio ao longo destes anos.

Agradeço ao anjo que tomou conta das minhas princesas quando eu precisei, dona Cleonice.

A todos os meus amigos que me deram suporte ao longo desta jornada.

À CAPES, CNPq e CRUB-UNCo pelo apoio financeiro.



# Resumo

Nesta tese apresentamos estratégias para resolver uma classe importante de problemas de tomada de decisões. Nosso propósito consiste em minimizar uma determinada função objetivo  $F(x)$ , onde  $F : \Re^n \rightarrow \Re$  e  $x$  é um ponto propriamente eficiente do problema multiobjetivo  $\min_{x \in \mathcal{X}} f(x)$ , com  $f(x) = (f_1(x), \dots, f_p(x))$ . Este problema é formulado como um problema em dois níveis, onde com um escalamento conveniente, podemos garantir pontos propriamente eficientes do problema vetorial do nível inferior, usando resultados de Geoffrion (1968). Este tipo particular de ponto é útil para a tomada de decisões ótimas sem negligenciar nenhum dos objetivos da função vetorial.

Também analisamos uma dificuldade presente nos métodos de Restauração Inexata na resolução de problemas com restrições de complementaridade e mostramos o artifício para contornar a complicação.

Para resolver o problema em dois níveis apresentamos estratégias de Restauração Inexata (RI) e as comparamos com a estratégia da reformulação, na qual se utilizam as condições de Karush-Kuhn-Tucker (KKT) do problema do segundo nível como restrições de um problema de programação não linear. Também apresentamos resultados numéricos que evidenciam que as estratégias de RI são melhores que a reformulação KKT.

# Abstract

In this thesis, we present strategies for solving an important class of problems in the decision making. The problem consists in minimising a determinate objective function  $F(x)$ , where  $F : \Re^n \rightarrow \Re$  and  $x$  is a properly efficient point of the multiobjective problem  $\min_{x \in \mathcal{X}} f(x)$ , with  $f(x) = (f_1(x), \dots, f_p(x))$ .

This problem is formulated as a bilevel problem, where with a convenient scaling, we can guarantee points properly efficient of the vectorial problem of the inferior level, cf. Geoffrion (1968). This particular type of point is useful for making optimum decisions without neglecting any of the objectives of the vectorial function.

We have also detected a difficulty of the Inexact Restoration methods in the resolution of problems with complementarity constraints and show an artifice for solving the complication.

To solve the bilevel problem, we present strategies of Inexact Restoration and compare them with the traditional strategy that uses the KKT conditions of the problem from the second level. We also present numerical experiments that evidence that the RI strategies are better than the traditional method.

# Lista de Figuras

2.1	Ilustração para o problema multiobjetivo do Exemplo 1. . . . .	24
2.2	Conjunto solução de pontos propriamente eficientes. . . . .	26
2.3	Pontos propriamente eficientes segundo Kuhn e Tucker mas não segundo Geoffrion. . . . .	33
2.4	Conjunto de pontos eficientes $\mathcal{X}^*$ correspondente ao problema dado em (2.33). . . . .	41
5.1	Conjunto $\mathcal{X}^*$ do problema 10. . . . .	73
5.2	Conjunto $\mathcal{X}^*$ do problema 12. . . . .	74
5.3	Conjunto $\mathcal{X}^*$ do problema 13. . . . .	75
5.4	Ilustração de $\mathcal{X}^*$ para o problema 20, com $p = 3$ . . . . .	79
6.1	Gráficos para problemas DTLZ sem restrições (Problemas 17, 18 e 19). . . . .	113
6.2	Gráficos para problemas DTLZ com restrições (Problemas 23, 24 e 25) . . . . .	113

# Lista de Tabelas

4.1	Tabela comparativa. Ponto intermediário . . . . .	64
6.1	Problemas sem restrições - Método 3 . . . . .	92
6.2	Problemas sem restrições - Método 1 . . . . .	93
6.3	Problemas sem restrições - Método 2 . . . . .	94
6.4	Problemas com restrições - Método 3 . . . . .	95
6.5	Problemas com restrições - Método 1 . . . . .	96
6.6	Problemas com restrições - Método 2 . . . . .	96
6.7	Família de problemas sem restrições - Método 3 . . . . .	97
6.8	Família de problemas sem restrições - Método 1 . . . . .	98
6.9	Família de problemas sem restrições - Método 2 . . . . .	98
6.10	Família de problemas com restrições - Método 3 . . . . .	99
6.11	Família de problemas com restrições - Método 1 . . . . .	100
6.12	Família de problemas com restrições - Método 2 . . . . .	101
6.13	Tabela comparativa dos problemas DTLZ1 e DTLZ2 sem restrições (Problemas 17 e 18). . . . .	104
6.14	Tabela comparativa do problema DTLZ3 sem restrições (problema 19). . .	105
6.15	Tabela comparativa do problema DTLZ1 com restrições (Problema 23). . .	106
6.16	Tabela comparativa do problema DTLZ2 com restrições (Problema 24). . .	107
6.17	Tabela comparativa do problema DTLZ3 com restrições (Problema 25). . .	108
6.18	Tabela comparativa separando por dimensão. Problemas sem restrições (17, 18 e 19). . . . .	109
6.19	Tabela comparativa agrupando as dimensões e separando pela inicialização adotada. Problemas sem restrições (17, 18 e 19). . . . .	109

6.20	Tabela comparativa separando por dimensão. Problemas com restrições (23, 24 e 25).	110
6.21	Tabela comparativa agrupando as dimensões e separando pela inicialização adotada. Problemas com restrições (23, 24 e 25).	110
6.22	Resultados por dimensão. Problemas com restrições (23, 24 e 25)	111
6.23	Resultados por ponto gerado. Problemas com restrições (Problemas 23, 24 e 25)	112

Notação	Significado
PNL	Programação não linear.
$\Omega$ ,	Domínio de função escalar.
$\mathcal{A}$	Domínio de função escalar.
$\mathcal{X}$ ,	Conjunto factível de um problema de PNL.
$\Upsilon$ ,	Conjunto factível de um problema de PNL.
$k, n, m, p$ ,	Números naturais.
$\mathbb{R}$ ,	Conjunto de números reais.
$\mathbb{R}^k$ ,	Espaço vetorial sobre $\mathbb{R}$ de dimensão $k$ .
$\mathbb{R}_+^k$ ,	$\{x \in \mathbb{R}^k \mid x_i \geq 0, \forall i = 1, \dots, k\}$ .
$\mathbb{R}_{++}^k$ ,	$\{x \in \mathbb{R}^k \mid x_i > 0, \forall i = 1, \dots, k\}$ .
$x = (x_1, \dots, x_n)$ ,	Vetor variável de $\mathbb{R}^n$ .
$\alpha = (\alpha_1, \dots, \alpha_p)$ ,	Vetor pesos de $\mathbb{R}_+^p$ .
$f = (f_1, \dots, f_p)$ ,	Função vetorial.
$f_i$ ,	Componente i-ésima de uma função vetorial.
$\nabla g$	Gradiente da função escalar $g$ , vetor coluna .
$\nabla^2 g$	Hessiana da função escalar $g$ .
$\ \cdot\ $	Norma Euclidiana em $\mathbb{R}^n$ .
$\ \cdot\ _\infty$	Norma Infinito em $\mathbb{R}^n$ .

# Sumário

<b>1</b>	<b>Introdução</b>	<b>17</b>
<b>2</b>	<b>Otimização Multiobjetivo</b>	<b>22</b>
2.1	Introdução . . . . .	22
2.2	Soluções eficientes ou Pareto-ótimas . . . . .	23
2.3	Pontos propriamente eficientes . . . . .	25
2.3.1	Definição, condições necessárias e suficientes e equivalência propriamente eficiente: Geoffrion e Kuhn-Tucker . . . . .	25
2.3.2	Conjuntos eficientes convexos . . . . .	40
<b>3</b>	<b>Restauração Inexata</b>	<b>45</b>
3.1	Introdução . . . . .	45
3.2	Descrição do modelo algorítmico . . . . .	46
3.2.1	Fase de restauração . . . . .	46
3.2.2	Aproximação linear . . . . .	47
3.2.3	Fase de minimização . . . . .	47
3.2.4	Função de mérito . . . . .	48
3.2.5	Algoritmo 1 . . . . .	49
3.2.6	Observações . . . . .	50
3.3	Hipóteses gerais e consequências . . . . .	52
3.4	Restauração Inexata para programação em dois níveis . . . . .	55
<b>4</b>	<b>Reformulação de um problema de PNL com restrições multiobjetivo</b>	<b>59</b>
4.1	Introdução . . . . .	59
4.2	Métodos Implementados . . . . .	61
4.2.1	Método 1: Restauração Inexata 1 . . . . .	61
4.2.2	Método 2: Restauração Inexata 2 . . . . .	61
4.2.3	Método 3: Reformulação KKT . . . . .	62
4.3	Algoritmo RI-PDNM . . . . .	63
<b>5</b>	<b>Problemas testados</b>	<b>68</b>
5.1	Introdução . . . . .	68
5.2	Problemas testados . . . . .	68
5.2.1	Grupo 1 sem restrições . . . . .	69
5.2.2	Grupo 1 com restrições . . . . .	71
5.2.3	Grupo 2 sem restrições . . . . .	75

5.2.4	Grupo 2 com restrições . . . . .	79
<b>6</b>	<b>Resultados dos experimentos</b>	<b>84</b>
6.1	Introdução . . . . .	84
6.2	RI sem perturbação ou RI com perturbação . . . . .	84
6.3	Resultados obtidos . . . . .	90
6.3.1	Testes preliminares . . . . .	91
6.3.2	Experimentos sistematizados . . . . .	102
6.3.3	Gráficos de desempenho dos experimentos sistematizados . . . . .	112
6.3.4	Comentários sobre os experimentos . . . . .	114
<b>7</b>	<b>Conclusões</b>	<b>115</b>



---

---

# Capítulo 1

## Introdução

Neste trabalho fazemos uma proposta para a resolução de um importante problema de decisão. A meta é encontrar uma decisão ótima para uma dada função objetivo principal, sujeita a atender vários objetivos secundários, mas essenciais e conflitantes. Escolhemos critérios de proximidade com um ponto pre-estabelecido, o qual, na prática, poderia ser uma tomada de decisão que é adotada atualmente por uma empresa, ou ainda uma tomada de decisão utópica, que gostariam de adotar mas não podem. Isto pode ser formulado da seguinte maneira:

$$\begin{aligned} \min_x \quad & F(x) \\ \text{s.a} \quad & x \in \mathcal{X}^*, \end{aligned} \tag{1.1}$$

onde  $F : \Omega \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ ,  $\mathcal{X}^*$  corresponde ao conjunto de pontos que são solução do problema multiobjetivo

$$\begin{aligned} \min_x \quad & (f_1(x), \dots, f_p(x)) \\ \text{s.a} \quad & x \in X, \end{aligned}$$

$f_i : \mathcal{A} \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ ,  $\mathcal{A} \subseteq \Omega$ ,  $\forall i = 1, \dots, p$ , com  $X = \{x \in \mathbb{R}^n \mid h(x) = 0, g(x) \leq 0\}$ ,  $h$  e  $g$  funções vetoriais e todas funções de classe  $C^1$  em seus domínios.

O formato deste problema não é de um problema de programação não linear clássico (ver [17], [27], [28], [29], [33]), dado que as restrições contêm um problema multiobjetivo (ver [9], [15], [24], [25]). Propomos reformular (1.1) como um problema de otimização em dois níveis (ver [11], [12]), reescrevendo o problema multiobjetivo com uma reformulação com pesos (ver [9], [16]). Impondo que os pesos sejam maiores que ou iguais a um valor positivo pre-determinado, estamos atendendo ao que se chama pontos Pareto próprios,

que permitem conseguir soluções estáveis no sentido de que uma pequena perturbação nos dados não leva a pontos indesejáveis, além de não negligenciar nenhum dos objetivos analisados (ver Geoffrion [19]).

Desta maneira, dado  $\varepsilon > 0$ , que indica o nível mínimo de importância atribuído a cada um dos objetivos, o problema reformulado é:

$$\begin{aligned}
 & \min_{\alpha, x} && F(x) \\
 & \text{s.a} && \sum_{i=1}^p \alpha_i = 1, \\
 & && \alpha_i \geq \varepsilon, \\
 & && x \in \operatorname{argmin} \left\{ \min_{x \in X} \sum_{i=1}^p \alpha_i f_i(x) \right\}.
 \end{aligned} \tag{1.2}$$

É preciso ressaltar que as restrições do problema de segundo nível em (1.2) não podem ser levadas ao primeiro nível, sob pena de modificar o problema original.

A estratégia proposta será a de Restauração Inexata (RI) como a dada em [2]. Esta abordagem consiste em aproveitar a estrutura do problema de otimização do segundo nível, fazendo uma busca exigente do mínimo na fase de restauração, e aproveitando os excelentes algoritmos para problemas de otimização que existem. Tal estrutura é importante neste caso, dadas as características deste problema, porque até pode obter-se o mínimo global, e logo aproveitar a reformulação KKT do segundo nível para a fase de otimalidade. Em contraposição, com esta estratégia temos a reformulação do problema geral como problema de otimização usando as condições KKT do problema de segundo nível [1], o qual descaracteriza totalmente o segundo nível, desconsiderando a estrutura do problema de otimização e reduzindo-o simplesmente a um problema variacional. Mostramos experimentos numéricos que avaliam a vantagem da nossa estratégia.

O problema (1.2) é claramente um problema em dois níveis no qual a função objetivo do primeiro nível depende somente da variável de decisão. Isto pode gerar circunstâncias indesejáveis, tanto para a estratégia de Restauração Inexata quanto para a estratégia de reformulação KKT do problema do segundo nível.

Na abordagem proposta, uma vez fornecidos os pesos e concluída a fase de restauração, tem início a fase de otimalidade. A reformulação KKT do problema do segundo nível obriga a inclusão de uma variável não negativa e uma restrição de complementaridade para cada restrição de desigualdade do problema do segundo nível. Se esta nova variável

for positiva na inicialização, o qual até pode ser desejável porque estaríamos em um ponto com complementaridade estrita, isto obriga a restrição correspondente a converter-se em restrição de igualdade. Ao ignorar a parte viável e interior dessa restrição, podem ser gerados pontos KKT do problema reformulado, mas candidatos extremamente ruins para a solução do problema original. Para ilustrar esta situação apresentamos o seguinte exemplo de um problema em dois níveis: Seja  $\epsilon > 0$  e considerar o problema

$$\begin{aligned}
 \min_{\alpha, x} \quad & x_1^2 + x_2^2 \\
 \text{s.a} \quad & \alpha_1 + \alpha_2 + \alpha_3 = 1 \\
 & \alpha_1, \alpha_2, \alpha_3 \geq \epsilon \\
 & x = \arg \min_x \quad \alpha_1 f_1(x) + \alpha_2 f_2(x) + \alpha_3 f_3(x) \\
 & \text{s. a} \quad x_1 + x_2 - 0.8 \leq 0, \\
 & \quad \quad \quad x_1, x_2 \geq 0,
 \end{aligned} \tag{1.3}$$

onde  $f_1(x) = x_1^2 + (x_2 - 1)^2$ ,  $f_2(x) = (x_1 - 1)^2 + x_2^2$  e  $f_3(x) = x_1^2 + x_2^2$ .

A solução do problema de segundo nível, parametrizado em  $\alpha$ , corresponde a um subconjunto próprio convexo e fechado contido no triângulo de vértices  $(0.8, 0)$ ,  $(0, 0.8)$  e  $(0, 0)$ , assumindo  $\epsilon \leq 1/3$ . Mais informações sobre soluções desta natureza são dadas no capítulo 2 desta tese, em que o problema do segundo nível é tratado como um problema multiobjetivo. Logo a solução do problema (1.3) corresponde ao ponto  $(\epsilon, \epsilon)$ .

Analizamos a resolução deste problema usando o método de RI. Escolhendo  $\epsilon \leq 0.1$ , agregando folga na restrição do problema de segundo nível e chamando de  $x_3$  a esta nova variável, o problema de segundo nível resulta:

$$\begin{aligned}
 \min_{x \in \mathbb{R}^3} \quad & \alpha_1 f_1(x) + \alpha_2 f_2(x) + \alpha_3 f_3(x) \\
 \text{s.a} \quad & x_1 + x_2 + x_3 - 0.8 = 0, \\
 & x_1, x_2, x_3 \geq 0.
 \end{aligned} \tag{1.4}$$

Se o ponto que é solução para este subproblema, que corresponde à fase de restauração do método de RI, é tal que a variável dual associada a  $x_3$  é estritamente maior que zero, pela complementaridade estrita das condições KKT deste último problema, deve-se cumprir que  $x_3 = 0$ . Assim, na fase de otimalidade, a minimização é feita sobre a reta  $x_1 + x_2 - 0.8 = 0$ . Pode-se conferir que a solução do subproblema da fase de otimalidade é o ponto obtido na fase de restauração.

Dessa forma, esse ponto é um ponto KKT do subproblema proposto da fase de otimização, mas não corresponde a um mínimo local do problema original. Mais ainda, podemos dizer que corresponde a um ponto extremamente ruim do ponto de vista da decisão original.

Solucionamos esta dificuldade com a estratégia algorítmica de resolução por Restauração Inexata em duas fases, agregando uma perturbação que converge para zero na reformulação das restrições de desigualdade do problema do nível inferior.

As experiências numéricas realizadas para resolver o problema (1.2) consistem em duas versões de nossa proposta de Restauração Inexata, usando na fase de otimalidade a função lagrangiana ou a função objetivo, que chamamos de Método 1 e Método 2, respectivamente e, a reformulação KKT, que chamamos de Método 3. Esta última consiste em minimizar a função objetivo do primeiro nível sujeita às condições KKT do problema do segundo nível.

Finalmente podemos destacar algumas das vantagens de nossa estratégia. Primeiro, como se considera a estrutura do problema de segundo nível, é possível achar um subconjunto particular de soluções do problema multiobjetivo e assim obter melhores resultados que com a reformulação KKT. Por outro lado, é possível trabalhar o problema de otimização vetorial com muitos objetivos, o que em geral não é feito pela complexidade que isto envolve e além disso, utilizar métodos de programação não linear na resolução de problemas multiobjetivo, diferentemente de outros trabalhos que utilizam métodos heurísticos para resolvê-los.

A seguir é detalhada a estrutura deste trabalho:

**Capítulo 2:** Neste capítulo são apresentados alguns conceitos básicos da Otimização Multiobjetivo. Definimos o conceito de solução para problemas vetoriais, os quais são conhecidos na literatura como pontos eficientes. Dentro do conjunto de pontos eficientes, diferenciamos o conjunto de pontos *propriamente eficientes*. Na literatura existem várias definições destes pontos, nesta tese nos focalizamos em duas: no sentido de Geoffrion e Kuhn-Tucker. Neste capítulo expomos e provamos alguns resultados teóricos relacionados com estes pontos em ambos os sentidos e demonstramos a equivalência existente entre eles. Também mostramos exemplos nos quais

há pontos que satisfazem uma definição de propriamente eficiente e a outra não. Finalmente apresentamos resultados teóricos os quais determinarão condições sob as quais o conjunto de pontos propriamente eficiente será um conjunto convexo, o que em geral não é verdade.

**Capítulo 3:** Aqui é tratado o método de Restauração Inexata para Programação Não Linear. Neste capítulo expomos alguns resultados teóricos referentes a este método e mostramos a adaptação do algoritmo para problemas de programação em dois níveis, que foi desenvolvido em [2].

**Capítulo 4:** Este capítulo corresponde aonexo entre os dois capítulos anteriores. A estrutura do problema que estamos interessados em resolver é inspirada nos problemas de programação em dois níveis, mas não possui originalmente esta estrutura. No capítulo 4, sob certas hipóteses, faremos a reformulação destes problemas como problemas de programação em dois níveis, que chamaremos de “*Problemas de Programação Bilevel-Multiobjetivo*”. Também serão detalhados os três métodos implementados e apresentaremos o algoritmo para problemas em dois níveis-multiobjetivo, segundo o Método 2.

**Capítulo 5:** Para comparar o desempenho dos métodos, fizemos testes computacionais. Para os experimentos preliminares, criamos problemas nos quais foram satisfeitas as hipóteses de convexidade e para testar a robustez dos métodos, problemas nos quais não era satisfeita tal hipótese. Também implementamos problemas sistematizados os quais retiramos da literatura. Neste capítulo listamos os problemas que foram resolvidos.

**Capítulo 6:** Neste capítulo retomamos a discussão surgida com o problema (1.3) e mostramos esta dificuldade em um dos problemas práticos testados. Também exibimos os resultados obtidos dos experimentos computacionais por meio de tabelas comparativas e gráficos, entre eles gráficos de perfil de desempenho.

**Capítulo 7:** Neste último capítulo são analisados os resultados obtidos dos experimentos numéricos e, apresentamos as conclusões.

---

---

## Capítulo 2

# Otimização Multiobjetivo

### 2.1 Introdução

Uma classe importante de problemas de otimização é a otimização multiobjetivo ou também conhecida como otimização multicritério. A primeira noção neste tipo de problema remonta a Edgeworth (1881) [14] e Pareto (1896) [34] na área da economia. Mais tarde, em 1951, Kuhn e Tucker [26] fizeram a primeira consideração matemática do tema. Desde aquela época a otimização multiobjetivo tornou-se um campo de pesquisa ativa.

Este tipo de problema surge quando é preciso tomar decisões na presença de múltiplos objetivos ou critérios. A otimização é feita sobre uma função vetorial e as soluções deste tipo de problemas são conhecidas na literatura como soluções não inferiores, eficientes, Pareto-ótimas (em homenagem a V. Pareto), ou, não dominadas. Neste trabalho adotaremos o nome de soluções eficientes ou Pareto-ótimas. A dificuldade que surge nestes problemas é que o ótimo de um dos objetivos dará prejuízo no outro, isto é, haverá um *conflito de interesses*. Assim, tem que se encontrar um estado de “*equilíbrio*”, no qual todos os objetivos fiquem satisfeitos em um grau aceitável.

O conceito de solução é diferente ao da otimização com um objetivo somente, ou no sentido usual, no qual dizemos que  $x^*$  é um ponto de mínimo da função  $f(x)$  em um conjunto  $\mathcal{B}$  se, e somente se,

$$f(x^*) \leq f(x), \quad \forall x \in \mathcal{B}.$$

Em geral, não existe uma solução que otimize (minimize ou maximize) simultaneamente todos os objetivos. Uma solução é considerada ótima, se não existe outra solução ad-

missível que melhore simultaneamente todos os objetivos; a melhoria num objetivo é alcançada à custa da degradação do valor de pelo menos um dos outros objetivos. Neste caso temos que usar uma ordem parcial, onde podemos dizer que  $v(x_1)$  domina  $v(x_2)$  se todos os elementos do vetor  $v(x_1)$  são maiores ou iguais que os elementos de  $v(x_2)$  e denotaremos  $v(x_1) \geq v(x_2)$ .

## 2.2 Soluções eficientes ou Pareto-ótimas

Sejam  $x \in \mathbb{R}^n$  o vetor de variáveis de decisão,  $g_i(x)$ ,  $i = 1, \dots, m$ , a  $i$ -ésima restrição com valores reais definida em  $\mathbb{R}^n$ . A região viável, ou espaço de decisão, será o conjunto

$$\mathcal{X} = \{x \in \Omega \mid g_i(x) \leq 0, i = 1, \dots, m\},$$

onde  $\Omega \subseteq \mathbb{R}^n$  representa o conjunto das restrições que não podem ser expressas na forma funcional  $g$ .

O *Problema Multiobjetivo* (**PM**) na forma geral é definido como

$$\min_{x \in \mathcal{X}} f(x)$$

onde  $f : \mathcal{A} \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^p$ . A imagem de  $f$ ,

$$\mathcal{F}(\mathcal{X}) = \{f(x) \mid x \in \mathcal{X}\}$$

é chamado o espaço objetivo.

**Definição 2.1** Um ponto  $x^* \in \mathcal{X}$  se diz eficiente se não existe outro  $x \in \mathcal{X}$  tal que  $f(x) \leq f(x^*)$ , isto é,  $f_j(x) \leq f_j(x^*) \forall j = 1, \dots, p$  cumprindo-se a desigualdade estrita para pelo menos um  $j$ .

Vamos chamar o conjunto de soluções eficientes  $\mathcal{X}^*$  e o conjunto imagem associado de  $\mathcal{F}^* = \{f(x) \mid x \in \mathcal{X}^*\}$ . Neste trabalho, estamos interessados somente no conjunto de soluções eficientes, isto é, em  $\mathcal{X}^*$ .

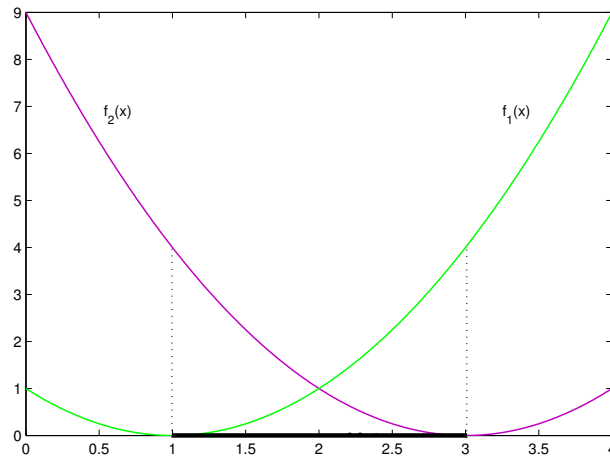
Intuitivamente, um ponto  $x^*$  é eficiente se qualquer outro  $x \in \mathcal{X}$  alternativo, tal que melhore o valor em alguma das funções objetivo (diminua) em relação ao valor em  $x^*$ , deve, ao mesmo tempo, aumentar o valor em alguma outra função objetivo.

Para ilustrar o conceito analisamos o problema a seguir:

**Exemplo 1:**

$$\begin{aligned} \min \quad & (f_1(x), f_2(x)) \\ \text{s. a} \quad & x \geq 0, \end{aligned}$$

onde  $f_1(x) = (x - 1)^2$  e  $f_2(x) = (x - 3)^2$ .



**Figura 2.1:** Ilustração para o problema multiobjetivo do Exemplo 1.

Neste exemplo  $\mathcal{X}^* = \{x \in \mathbb{R} \mid 1 \leq x \leq 3\}$  e  $\mathcal{F}^* = \{y \in \mathbb{R} \mid 0 \leq y \leq 4\}$ . Se  $x \notin [1; 3]$  sempre é possível achar  $z \in \mathcal{X}$  onde ambas as funções objetivo diminuam.

A seguir definimos a notação que usaremos para vetores neste trabalho.

Sejam  $a$ ,  $b$  vetores: (i)  $a = b$  se, e somente se,  $a_i = b_i \forall i$ , isto é, os vetores são iguais em cada uma das suas componentes, (ii)  $a \leq b$  se, e somente se,  $a_i \leq b_i \forall i$ , isto é, verifica-se a desigualdade em cada uma das componentes dos vetores e, (iii)  $a < b$  se, e somente se,  $a_i < b_i \forall i$ , isto é, verifica-se a desigualdade estrita em cada uma das componentes dos vetores.



## 2.3 Pontos propriamente eficientes

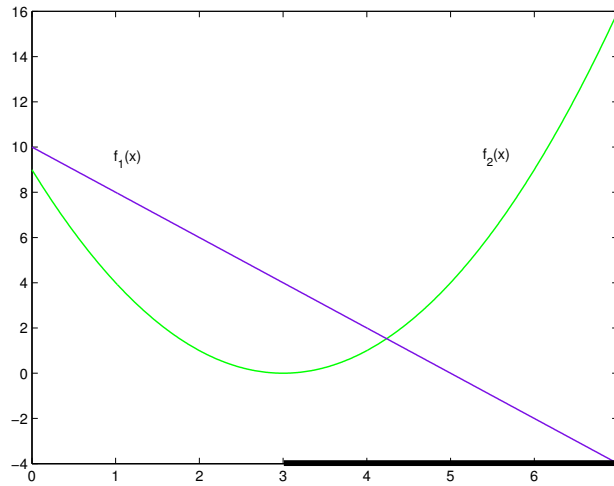
### 2.3.1 Definição, condições necessárias e suficientes e equivalência propriamente eficiente: Geoffrion e Kuhn-Tucker

Da seção anterior, temos que em pontos eficientes o valor de uma das funções objetivo aumenta à custa da deterioração de pelo menos um dos outros objetivos. Esta relação pode ser medida pela taxa de decrescimento em um dos objetivos, suponhamos  $f_i$ , em relação ao crescimento de outro  $f_j$ . Em algumas situações esta taxa pode não ser limitada, isto é, pode ser feita arbitrariamente grande. Os primeiros que observaram esta propriedade foram Kuhn e Tucker ((1951), [26]), mais tarde Geoffrion ((1968),[19]) modificou a caracterização desses pontos. As duas definições não incluem os mesmos pontos, mas sob certas hipóteses elas são equivalentes, como aparece em alguns trabalhos, entre eles [15] no qual nos baseamos para esta seção. Borwein ((1977)[7],[15]) e Benson ((1979)[3],[15]) também analisaram essa questão, conforme detalharemos nesta seção.

**Definição 2.2** (Geoffrion) Dizemos que  $x^* \in \mathcal{X}^*$  é propriamente eficiente se existe um escalar  $M > 0$  tal que, para cada  $i = 1, \dots, p$ , e cada  $x \in \mathcal{X}$  tal que  $f_i(x) < f_i(x^*)$ , existe pelo menos um  $j \neq i$  com  $f_j(x) > f_j(x^*)$  e  $\frac{f_i(x^*) - f_i(x)}{f_j(x) - f_j(x^*)} \leq M$ . Uma solução eficiente que não é propriamente eficiente é denominada solução eficiente imprópria.

Notar que, da definição, a constante não negativa  $M$  depende de cada ponto propriamente eficiente, ou seja, temos  $M = M(x^*)$ . Por outro lado, podemos dizer que um ponto é *propriamente eficiente* no sentido de Geoffrion, quando a taxa que mede o decrescimento-crescimento das funções objetivo é limitada.

Analisamos o exemplo da seguinte figura:



**Figura 2.2:** Conjunto solução de pontos propriamente eficientes.

Neste exemplo, o problema

$$\begin{aligned} \min \quad & (f_1(x), f_2(x)) \\ \text{s. a} \quad & x \geq 0 \end{aligned} \tag{2.1}$$

com  $f_1(x) = 10 - 2x$  e  $f_2(x) = (x - 3)^2$ , tem como solução  $\mathcal{X}^* = \{x \in \mathbb{R} \mid x \geq 3\}$ . O único ponto impróprio de  $\mathcal{X}^*$  é  $x = 3$  pois, se tomarmos  $z = x + h$ , com  $h > 0$ ,  $f_1(z) < f_1(x)$ ,  $f_2(z) > f_2(x)$  mas

$$\frac{f_1(x) - f_1(z)}{f_2(z) - f_2(x)} = \frac{\frac{f_1(x) - f_1(z)}{h}}{\frac{f_2(z) - f_2(x)}{h}} \rightarrow \infty \quad \text{se } h \rightarrow 0^+.$$

Portanto,  $x = 3$  é impróprio. Analisemos os demais pontos de  $\mathcal{X}^*$ . Vamos tomar  $x \in (3, \infty)$  arbitrário e  $h > 0$ , logo  $x + h \in \mathcal{X}^*$ ,  $f_2(x) < f_2(x + h)$ ,  $f_1(x) > f_1(x + h)$  e

$$\frac{f_1(x) - f_1(x + h)}{f_2(x + h) - f_2(x)} = \frac{2}{h + 2(x - 3)} < \frac{1}{x - 3}.$$

Assim escolhendo  $M \geq \frac{1}{x - 3}$ , temos que  $x \in (3, \infty)$  é propriamente eficiente. ■

Na Figura 2.2 ilustramos o problema (2.1), no qual o conjunto de pontos eficientes possui um ponto em particular ( $x = 3$ ) para o qual o problema apresenta uma instabilidade. Para a função  $f_2(x) = (x - 3)^2$ , em  $3 + h$ ,  $h > 0$ , ocorre um aumento de segunda ordem, enquanto para a função  $f_1(x) = 10 - 2x$ , também em  $3 + h$ ,  $h > 0$ , ocorre uma redução de primeira ordem. Isto implica que localmente temos uma subida mais lenta que a descida, o que gera uma instabilidade no problema multiobjetivo.

Para o caso de problemas lineares, qualquer ponto do conjunto  $\mathcal{X}^*$  é propriamente eficiente, ver [23].

Uma estratégia utilizada para achar pontos eficientes é transformar o problema multiobjetivo em um problema de otimização escalar apropriado (ver [9],[16], [21], [22]). Existem varias opções para tal fim, mas neste trabalho estamos interessados no “*problema com pesos*”. Seja

$$W = \{w \mid w \in \mathbb{R}^p, w_j \geq 0 \text{ e } w \neq 0\}.$$

Definimos para algum  $w \in W$ , o problema  $P(w)$  :

$$\min_{x \in \mathcal{X}} \sum_{j=1}^p w_j f_j(x). \quad (2.2)$$

A seguir são apresentados teoremas que, sob certas hipóteses, determinam a relação existente entre pontos propriamente eficientes e as soluções do problema escalar com pesos.

**Teorema 2.1** *Sejam  $\lambda_i > 0$  ( $i = 1, \dots, p$ ) fixos. Se  $x^*$  é uma solução de  $P(\lambda)$ , então  $x^*$  é propriamente eficiente para o problema (PM).*

**Prova:** Para provar que  $x^*$  é propriamente eficiente, primeiro devemos provar que  $x^*$  é eficiente. Suponhamos, pelo contrário, que  $x^*$  não é eficiente, então  $\exists z \in \mathcal{X}$  tal que

$$f_i(z) \leq f_i(x^*) \quad \forall i = 1, \dots, p \text{ e } f_j(z) < f_j(x^*) \text{ para algum } j \text{ com } 1 \leq j \leq p.$$

Como, pela hipótese,  $\lambda_i > 0 \quad \forall i = 1, \dots, p$ , então  $\lambda_i f_i(z) \leq \lambda_i f_i(x^*) \quad \forall i = 1, \dots, p$  e  $\lambda_j f_j(z) < \lambda_j f_j(x^*)$  para algum  $j$  com  $1 \leq j \leq p$ . Logo, somando para  $i = 1, \dots, p$ , obtemos a seguinte desigualdade

$$\sum_{i=1}^p \lambda_i f_i(z) < \sum_{i=1}^p \lambda_i f_i(x^*).$$

Mas isto contradiz o fato de  $x^*$  ser uma solução ótima de  $P(\lambda)$ . Portanto  $x^*$  é solução eficiente.

A prova da segunda condição, de que  $x^*$  é um ponto propriamente eficiente, também é feita por absurdo. Suponhamos que  $\forall M > 0, \exists i$  com  $1 \leq i \leq p$  e  $\exists x \in \mathcal{X}$  tais que

$$\frac{f_i(x^*) - f_i(x)}{f_j(x) - f_j(x^*)} > M, \quad \forall j \neq i \text{ tal que } f_i(x) < f_i(x^*) \text{ e } f_j(x^*) < f_j(x).$$

Operando algebricamente a desigualdade anterior obtemos

$$f_i(x^*) - f_i(x) > M (f_j(x) - f_j(x^*)). \quad (2.3)$$

Como esta desigualdade é válida  $\forall M > 0$ , em particular vale se escolhermos

$$M = (p-1) \max_{r,s} \left\{ \frac{\lambda_r}{\lambda_s} \right\},$$

assumindo que  $p \geq 2$ . Substituindo este valor em (2.3) obtemos

$$f_i(x^*) - f_i(x) > (p-1) \max_{r,s} \left\{ \frac{\lambda_r}{\lambda_s} \right\} (f_j(x) - f_j(x^*)). \quad (2.4)$$

Por outro lado,

$$(p-1) \max_{r,s} \left\{ \frac{\lambda_r}{\lambda_s} \right\} (f_j(x) - f_j(x^*)) \geq \frac{(p-1)\lambda_j}{\lambda_i} (f_j(x) - f_j(x^*)) \quad \forall j \neq i. \quad (2.5)$$

De (2.4) e (2.5) temos

$$f_i(x^*) - f_i(x) > \frac{(p-1)\lambda_j}{\lambda_i} (f_j(x) - f_j(x^*)) \quad \forall j \neq i. \quad (2.6)$$

Multiplicando por  $\lambda_i/(p-1)$  em ambos os lados da desigualdade anterior (2.6) e somando sobre  $j \neq i$  obtemos

$$\lambda_i (f_i(x^*) - f_i(x)) > \sum_{j \neq i} \lambda_j (f_j(x) - f_j(x^*)). \quad (2.7)$$

Logo, agrupando convenientemente, temos

$$\sum_{j=1}^p \lambda_j f_j(x^*) > \sum_{j=1}^p \lambda_j f_j(x). \quad (2.8)$$

Mas isto contradiz o fato de  $x^*$  ser um ponto ótimo de  $P(\lambda)$ . Assim fica provado o teorema. ■

**Corolário 2.1** *Seja  $\lambda \in \mathbb{R}_+^p$  com  $\lambda_1 + \dots + \lambda_p = 1$ . Se  $x^*$  é a única solução do problema  $P(\lambda)$ , então  $x^*$  é um ponto eficiente para o problema  $(\mathbf{PM})$ .*

**Prova:** A prova é feita como a primeira parte do teorema anterior. Suponhamos que  $x^*$  não é eficiente, então  $\exists z \in \mathcal{X}$  tal que

$$f_i(z) \leq f_i(x^*) \quad \forall i = 1, \dots, p \quad \text{e} \quad f_j(z) < f_j(x^*) \quad \text{para algum } j \text{ com } 1 \leq j \leq p.$$

Então, em particular para este  $z$ , temos que  $f(z) \leq f(x^*)$ .

Como, pela hipótese,  $\lambda_i \geq 0 \quad \forall i = 1, \dots, p$ , então  $\lambda_i f_i(z) \leq \lambda_i f_i(x^*) \quad \forall i = 1, \dots, p$ . Logo, somando para  $i = 1, \dots, p$ , obtemos a seguinte desigualdade

$$\sum_{i=1}^p \lambda_i f_i(z) \leq \sum_{i=1}^p \lambda_i f_i(x^*).$$

Mas isto contradiz o fato de  $x^*$  ser a única solução ótima de  $P(\lambda)$ . Portanto  $x^*$  é solução eficiente do problema  $(\mathbf{PM})$ . ■

**Teorema 2.2** *Seja  $\mathcal{X}$  um conjunto convexo, e sejam  $f_i$  convexas em  $\mathcal{X}$ . Então  $x^*$  é propriamente eficiente para  $(\mathbf{PM})$  se e somente se  $x^*$  é ótimo para  $P(\lambda)$  para algum  $\lambda$  com componentes estritamente positivas.*

**Prova:** Suponhamos que  $x^*$  é propriamente eficiente para  $(\mathbf{PM})$ , então  $\exists M > 0$  tal que  $\forall i = 1, \dots, p$

$$\frac{f_i(x) - f_i(x^*)}{f_j(x^*) - f_j(x)} \leq M$$

para algum  $j \neq i$  com  $f_j(x^*) < f_j(x)$  quando  $f_i(x) < f_i(x^*)$  e  $x \in \mathcal{X}$ .

Para cada  $i = 1, \dots, p$  seja  $F^i(x) : \mathcal{X} \rightarrow \mathbb{R}^p$ , a função vetorial cujas componentes são definidas como segue:

$$\begin{aligned} F_i^i(x) &= f_i(x) - f_i(x^*) \\ F_j^i(x) &= f_i(x) - f_i(x^*) + M(f_j(x) - f_j(x^*)) \quad \forall j \neq i. \end{aligned}$$

Como  $x^*$  é propriamente eficiente,  $\nexists x \in \mathcal{X}$  tal que  $F^i(x) < 0 \quad \forall i = 1, \dots, p$ , então pelos teoremas fundamentais para funções convexas ([28], página 63),  $\forall i = 1, \dots, p, \exists w^i \in \mathbb{R}^p$  com  $w^i \geq 0$  e  $w^i \neq 0$  tal que

$$w_i^i F_i^i(x) + \sum_{\substack{j=1 \\ j \neq i}}^p w_j^i F_j^i(x) \geq 0. \quad (2.9)$$

Por outro lado  $\sum_{k=1}^p w_k^i > 0$ , então multiplicando ambos os lados da desigualdade (2.9)

por  $1/\sum_{k=1}^p w_k^i$ , temos

$$\frac{w_i^i}{\sum_{k=1}^p w_k^i} F_i^i(x) + \sum_{\substack{j=1 \\ j \neq i}}^p \frac{w_j^i}{\sum_{k=1}^p w_k^i} F_j^i(x) \geq 0. \quad (2.10)$$

Chamando  $a_r^s = w_r^s / \sum_{k=1}^p w_k^s \quad \forall r, s$  e substituindo  $F^i(x)$  por sua expressão, obtemos:

$$a_i^i f_i(x) - a_i^i f_i(x^*) + \sum_{\substack{j=1 \\ j \neq i}}^p a_j^i [f_i(x) - f_i(x^*) + M(f_j(x) - f_j(x^*))] \geq 0. \quad (2.11)$$

Equivalentemente

$$a_i^i f_i(x) + \sum_{\substack{j=1 \\ j \neq i}}^p a_j^i [f_i(x) + M f_j(x)] \geq a_i^i f_i(x^*) + \sum_{\substack{j=1 \\ j \neq i}}^p a_j^i [f_i(x^*) + M f_j(x^*)]. \quad (2.12)$$

Operando algebricamente a desigualdade anterior obtemos

$$\sum_{j=1}^p a_j^i f_i(x) + M \sum_{\substack{j=1 \\ j \neq i}}^p a_j^i f_j(x) \geq \sum_{j=1}^p a_j^i f_i(x^*) + M \sum_{\substack{j=1 \\ j \neq i}}^p a_j^i f_j(x^*). \quad (2.13)$$

Observemos que  $\sum_{j=1}^p a_j^i = 1$ , então (2.13) fica

$$f_i(x) + M \sum_{\substack{j=1 \\ j \neq i}}^p a_j^i f_j(x) \geq f_i(x^*) + M \sum_{\substack{j=1 \\ j \neq i}}^p a_j^i f_j(x^*). \quad (2.14)$$

Somando para  $i = 1, \dots, p$  em (2.14) temos

$$\sum_{i=1}^p [f_i(x) + M \sum_{\substack{j=1 \\ j \neq i}}^p a_j^i f_j(x)] \geq \sum_{i=1}^p [f_i(x^*) + M \sum_{\substack{j=1 \\ j \neq i}}^p a_j^i f_j(x^*)]. \quad (2.15)$$

Reagrupando convenientemente

$$\sum_{j=1}^p [1 + M \sum_{\substack{j=1 \\ j \neq i}}^p a_j^i] f_j(x) \geq \sum_{j=1}^p [1 + M \sum_{\substack{j=1 \\ j \neq i}}^p a_j^i] f_j(x^*). \quad (2.16)$$

Agora definimos  $\lambda_j = 1 + M \sum_{i \neq j} a_j^i \forall j = 1, \dots, p$ , observando que  $\lambda_j > 0 \forall j$ . Substituindo em (2.16) obtemos

$$\sum_{j=1}^p \lambda_j f_j(x) \geq \sum_{j=1}^p \lambda_j f_j(x^*), \quad (2.17)$$

o que nos diz que  $x^*$  é solução ótima do problema  $P(\lambda)$  com  $\lambda > 0$  e é justamente o que estávamos procurando. A recíproca do teorema foi provada no *Teorema 2.1*. ■

No começo da seção 2.2 definimos o conjunto viável  $\mathcal{X}$  das restrições como:

$$\mathcal{X} = \{x \in \mathbb{R}^n \mid (g_1(x), \dots, g_m(x)) \leq 0\}.$$

Definimos agora o *conjunto dos índices das restrições ativas em  $x^*$*  como

$$\mathcal{J}(x^*) = \{j \in \{1, \dots, m\} \mid g_j(x^*) = 0\}.$$

Para a próxima definição de pontos propriamente eficientes, vamos supor que as funções objetivo  $f_i$ ,  $\forall i = 1, \dots, p$  e as funções  $g_j$ ,  $\forall j = 1, \dots, m$ , são continuamente diferenciáveis. Lembrando, o problema de programação multiobjetivo tem a seguinte estrutura:

$$\begin{aligned} \min_x \quad & f(x) \\ \text{s. a} \quad & g(x) \leq 0, \end{aligned}$$

com  $f : \mathbb{R}^n \rightarrow \mathbb{R}^p$  e  $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ .

Originalmente os pontos propriamente eficientes foram definidos da seguinte maneira:

**Definição 2.3** (*Kuhn-Tucker*). *Uma solução viável  $x^* \in \mathcal{X}$  é propriamente eficiente (no sentido de Kuhn e Tucker), se é eficiente e se não existe  $d \in \mathbb{R}^n$  tal que:*

$$\langle \nabla f_k(x^*), d \rangle \leq 0 \quad \forall k = 1, \dots, p, \quad (2.18)$$

$$\langle \nabla f_i(x^*), d \rangle < 0 \quad \text{para algum } i \in \{1, \dots, p\}, \quad (2.19)$$

$$\langle \nabla g_j(x^*), d \rangle \leq 0 \quad \forall j \in \mathcal{J}(x^*). \quad (2.20)$$

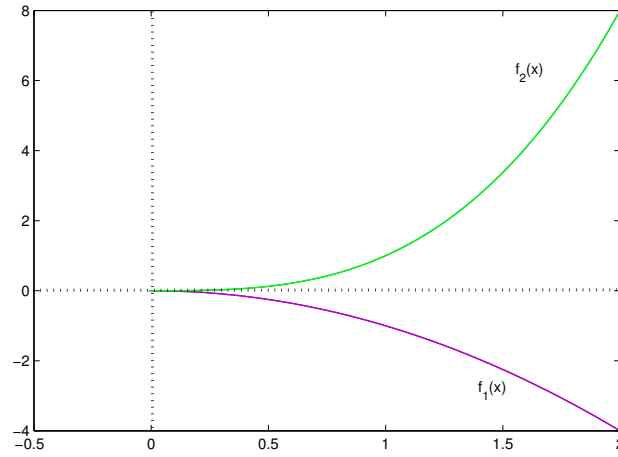
Voltando ao problema (2.1), podemos ver que o ponto  $x = 3$  também não é propriamente eficiente no sentido de Kuhn e Tucker pois,  $\nabla f_1(3) = -1$ ,  $\nabla f_2(3) = 0$  e  $\nabla g(3) = -1$ . Agora qualquer  $d > 0$  satisfaz as equações (2.18)-(2.20). Assim  $x = 3$ , não é propriamente eficiente nem no sentido de Geoffrion nem no sentido de Kuhn e Tucker.

Em geral estas duas definições não são equivalentes. A seguir apresentamos um exemplo que evidencia isto.

**Exemplo:** (Geoffrion (1968))

$$\begin{aligned} \min_{x \in \mathbb{R}} \quad & (-x^2, x^3) \\ \text{s. a} \quad & x \geq 0 \end{aligned}$$





**Figura 2.3:** Pontos propriamente eficientes segundo Kuhn e Tucker mas não segundo Geoffrion.

Os pontos eficientes para este problema correspondem aos números reais não negativos, isto é,  $\mathcal{X}^* = \{x \in \mathbb{R} \mid x \geq 0\}$ . O ponto  $x = 0$  é propriamente eficiente no sentido de Kuhn e Tucker, mas não no sentido de Geoffrion.

Chamamos  $f_1(x) = -x^2$ ,  $f_2(x) = x^3$  e  $g(x) = -x$ . Para provar que é satisfeita a Definição 2.3 calculamos os seguintes gradientes:

$$\begin{aligned} \nabla f_1(x) &= -2x, & \nabla f_1(0) &= 0, \\ \nabla f_2(x) &= 3x^2, & \nabla f_2(0) &= 0, \\ \nabla g(x) &= -1, & \nabla g(0) &= -1. \end{aligned}$$

Como  $\langle \nabla f_i(0), d \rangle = 0$ ,  $\forall i = 1, 2$ , e  $\forall d \in \mathbb{R}$ , temos que não existe  $d \in \mathbb{R}$  tal que as equações (2.18)-(2.20) sejam satisfeitas. Logo  $x = 0$  é um ponto propriamente eficiente no sentido de Kuhn e Tucker.

Analisemos agora a eficiência no sentido de Geoffrion. Seja  $h > 0$  e tomemos  $z = 0 + h$  então  $f_1(z) < f_1(0)$  e  $f_2(z) > f_2(0)$  portanto é eficiente, logo

$$\frac{f_1(0) - f_1(z)}{f_2(z) - f_2(0)} = \frac{h^2}{h^3} = \frac{1}{h} \rightarrow \infty \quad \text{se } h \rightarrow 0^+.$$

Portanto, em  $x = 0$  não é satisfeita a condição de ser limitada a taxa de decrescimento-crescimento dos objetivos. Logo o problema não é propriamente eficiente no sentido de Geoffrion.

Este exemplo apresenta uma diferença de ordem de variação perto do ponto propriamente eficiente  $(x^*)$ , isto é  $-x^2$  decresce mais rapidamente do que  $x^3$  cresce, isto gera instabilidade no ponto  $x^*$ . Por outro lado, a deterioração da função que diminui o seu valor não pode ser controlada com a melhoria da função que aumenta o seu valor. Pelo contrário, quando um ponto é propriamente eficiente no sentido de Geoffrion, a deterioração da função que diminui o seu valor pode ser controlada com o aumento da outra função. Assim, podemos dizer que a caracterização dada por Geoffrion para pontos propriamente eficientes é claramente um critério que tem melhor propriedade que a caracterização de Kuhn e Tucker.

Sob certas hipóteses, estas duas definições contêm os mesmos pontos. A seguir apresentamos a relação existente em ambos os sentidos, para isto é preciso supor uma condição de qualificação apropriada. O significado geométrico desta condição é que o conjunto viável  $\mathcal{X}$  tem uma descrição local em torno da solução viável  $x^*$  como um conjunto de curvas diferenciáveis partindo desse ponto: Toda direção viável  $d$ , pode ser escrita como o vetor velocidade a uma curva viável começando de  $x^*$ .

**Definição 2.4** *Um problema multiobjetivo diferenciável satisfaz a condição de qualificação KT (Kuhn-Tucker) em  $x^* \in \mathcal{X}$  se, para qualquer  $d \in \mathbb{R}^n$  tal que  $\langle \nabla g_j(x^*), d \rangle \leq 0 \forall j \in \mathcal{J}(x^*)$ , existem um número real  $\bar{t} > 0$ , uma função  $\theta : [0, \bar{t}] \rightarrow \mathbb{R}^n$ , e  $\alpha > 0$  tais que  $\theta(0) = x^*$ ,  $g(\theta(t)) \leq 0 \forall t \in [0, \bar{t}]$  e  $\theta'(0) = \alpha d$ .*

Vamos comparar esta condição com a condição de qualificação de Abadie em um ponto  $x^*$  ( $\text{CQA}(x^*)$ ) para problemas de programação não linear. Para isto precisamos de algumas definições prévias.

Consideremos o problema de programação não linear (PNL) na forma padrão:

$$\begin{aligned} \min_x \quad & f(x) \\ \text{s. a} \quad & h_i(x) = 0 \quad \forall i = 1, \dots, q, \\ & g_j(x) \leq 0 \quad \forall j = 1, \dots, m, \end{aligned}$$

com  $f, h_i, g_j : \mathbb{R}^n \rightarrow \mathbb{R} \forall i = 1, \dots, q, j = 1, \dots, m$  funções de classe  $C^1(\mathbb{R})$ . Chamaremos de  $K$  o conjunto viável do problema de PNL, isto é:

$$K := \{x \in \mathbb{R}^n \mid h_i(x) = 0 \ (i = 1, \dots, q), \ g_j(x) \leq 0 \ (j = 1, \dots, m)\}.$$

**Definição 2.5** *Seja  $\bar{x} \in K$ . Chamaremos o cone tangente para o PNL no ponto  $\bar{x}$  ao seguinte conjunto de pontos que denotaremos por  $\mathcal{T}_K(\bar{x})$ :*

$$\begin{aligned}\mathcal{T}_K(\bar{x}) &= \{d \in \mathbb{R}^n \mid \exists \{x^k\} \subset K, \exists t_k \rightarrow 0_+ : x^k \rightarrow \bar{x}, \frac{x^k - \bar{x}}{t_k} \rightarrow d\} \\ &= \{d \in \mathbb{R}^n \mid \exists \{d^k\}, d^k \rightarrow d, \exists t_k \rightarrow 0_+ : \bar{x} + t_k d^k \in K \forall k\}.\end{aligned}$$

*Chamaremos o cone linearizado para o PNL no ponto  $\bar{x}$  ao seguinte conjunto de pontos que denotaremos por  $\mathcal{L}_K(\bar{x})$*

$$\mathcal{L}_K(\bar{x}) = \{d \in \mathbb{R}^n \mid \nabla h_i(\bar{x})^T d = 0 \ (i = 1, \dots, q), \ \nabla g_j(\bar{x})^T d \leq 0 \ (j = 1, \dots, m)\}.$$

Das definições anteriores temos que

$$\mathcal{T}_K(x) \subseteq \mathcal{L}_K(x) \ \forall x \in K,$$

e a igualdade entre os cones é conhecida como a *condição de qualificação de Abadie*, enunciada a seguir:

**Definição 2.6** *Se diz que um ponto  $x^*$  satisfaz a condição de qualificação de Abadie (CQA( $x^*$ )) se*

$$\mathcal{T}_K(x^*) = \mathcal{L}_K(x^*).$$

Observar que a definição da condição de qualificação dada por KT em um ponto  $x^*$  é a versão de CQA( $x^*$ ) para curvas. Neste trabalho vamos usar a versão de KT porque não mudam os resultados e na literatura de referencia é usada esta última.

A seguir apresentamos a primeira relação na definição de pontos propriamente eficientes em ambos sentidos:

**Teorema 2.3** *(Geoffrion (1968)) Se um problema multiobjetivo diferenciável satisfaz a condição de qualificação KT em  $x^*$  e  $x^*$  é propriamente eficiente no sentido de Geoffrion (cf. Def. 2.2), então é propriamente eficiente no sentido de Kuhn e Tucker (cf. Def. 2.3).*

**Prova:** A demonstração será feita por absurdo. Seja  $x^*$  uma solução propriamente eficiente no sentido de Geoffrion, na qual é satisfeita a condição de qualificação KT, mas que

não é propriamente eficiente no sentido de Kuhn e Tucker. Então existe  $d \in \Re^n$  tal que

$$\langle \nabla f_s(x^*), d \rangle \leq 0 \quad \forall s = 1, \dots, p \quad (2.21)$$

$$\langle \nabla f_r(x^*), d \rangle < 0 \quad \text{para algum } r \in \{1, \dots, p\}, \quad (2.22)$$

$$\langle \nabla g_j(x^*), d \rangle \leq 0 \quad \forall j \in \mathcal{J}(x^*) \quad (2.23)$$

Como  $x^*$  satisfaz a condição de qualificação KT, existe  $\theta(t)$  função diferenciável tal que  $\theta(0) = x^*$  e  $\theta'(0) = d$ . Por outro lado  $x^*$  cumpre  $\langle \nabla f_r(x^*), d \rangle < 0$  para algum  $r \in \{1, \dots, p\}$  e, é um ponto eficiente. Então existem  $\{t_k\} \subset \Re_+$ ,  $t_k \rightarrow 0$  e um índice  $l$  tal que  $f_l(\theta(t_k)) \geq f_l(x^*)$  e  $f_r(\theta(t_k)) < f_r(x^*)$ . Fazendo a expansão de Taylor de  $f_l(x)$  em  $\theta(t_k)$  temos

$$f_l(\theta(t_k)) - f_l(x^*) = t_k \langle \nabla f_l(x^*), \theta'(t_k) \rangle + o(t_k),$$

logo

$$0 \leq \frac{f_l(\theta(t_k)) - f_l(x^*)}{t_k} = \langle \nabla f_l(x^*), \theta'(t_k) \rangle + \frac{o(t_k)}{t_k}.$$

Tomando limite temos

$$\langle \nabla f_l(x^*), d \rangle \geq 0.$$

Como, por hipótese,

$$\langle \nabla f_l(x^*), d \rangle \leq 0,$$

concluimos

$$\langle \nabla f_l(x^*), d \rangle = 0.$$

Continuando,

$$\frac{f_r(x^*) - f_r(\theta(t_k))}{f_l(\theta(t_k)) - f_l(x^*)} = \frac{-\langle \nabla f_r(x^*), \theta'(t_k) \rangle + \frac{o(t_k)}{t_k}}{\langle \nabla f_l(x^*), \theta'(t_k) \rangle + \frac{o(t_k)}{t_k}} \rightarrow \infty,$$

portanto  $x^*$  não é propriamente eficiente no sentido de Geoffrion, mas isto contradiz a hipótese. Com isto, fica provado o teorema. ■

Para provar a recíproca, precisamos de alguns resultados prévios. O primeiro teorema que apresentaremos para tal fim não será provado, mas a sua demonstração pode ser encontrada em [28], página 29.

**Teorema 2.4** (*Teorema alternativo de Tucker*) *Sejam  $B$ ,  $C$  e  $D$  matrizes  $p \times n$ ,  $l \times n$  e  $q \times n$  respectivamente. Então, ou o sistema de inequações*

$$Bx \leq 0 \quad e \quad Bx \neq 0, \quad Cx \leq 0, \quad Dx = 0 \quad (2.24)$$

*possui uma solução  $x \in \mathbb{R}^n$ , ou o sistema de equações*

$$B^T y^1 + C^T y^2 + D^T y^3 = 0, \quad y^1 > 0, \quad y^2 \geq 0 \quad (2.25)$$

*possui uma solução  $y^1 \in \mathbb{R}^p$ ,  $y^2 \in \mathbb{R}^l$ ,  $y^3 \in \mathbb{R}^q$ , mas nunca cumprem-se simultaneamente os dois.*

Com este teorema poderemos provar o seguinte resultado que estabelece as condições necessárias para pontos propriamente eficientes no sentido de Kuhn e Tucker.

**Teorema 2.5** *Se  $x^*$  é um ponto propriamente eficiente no sentido de Kuhn e Tucker, então existem  $\hat{\lambda} \in \mathbb{R}^p$  e  $\hat{\mu} \in \mathbb{R}^m$  tais que*

$$\sum_{k=1}^p \hat{\lambda}_k \nabla f_k(x^*) + \sum_{j=1}^m \hat{\mu}_j \nabla g_j(x^*) = 0, \quad (2.26)$$

$$\sum_{j=1}^m \hat{\mu}_j g_j(x^*) = 0, \quad (2.27)$$

$$\hat{\lambda} > 0, \quad (2.28)$$

$$\hat{\mu} \geq 0. \quad (2.29)$$

**Prova:** Como  $x^*$  é propriamente eficiente no sentido de Kuhn e Tucker (cf. Def. 2.3) não existe  $d \in \mathbb{R}^n$  tal que

$$\langle \nabla f_k(x^*), d \rangle \leq 0 \quad \forall k = 1, \dots, p, \quad (2.30)$$

$$\langle \nabla f_i(x^*), d \rangle < 0 \quad \text{para algum } i \in \{1, \dots, p\}, \quad (2.31)$$

$$\langle \nabla g_j(x^*), d \rangle \leq 0 \quad \forall j \in \mathcal{J}(x^*). \quad (2.32)$$

Logo, definindo as seguintes matrizes

$$\begin{aligned} B &= (\nabla f_k(x^*))_{k=1, \dots, p}, \\ C &= (\nabla g_j(x^*))_{j \in \mathcal{J}(x^*)}, \\ D &= 0, \end{aligned}$$

e aplicando o Teorema 2.4 temos que existe  $y^1 := \hat{\lambda}$ ,  $y^2 := \hat{\mu}$  e  $y^3$  com  $\hat{\lambda}_k > 0 \quad \forall k = 1, \dots, p$ ,  $\hat{\mu}_j \geq 0 \quad \forall j \in \mathcal{J}(x^*)$  satisfazendo

$$\sum_{k=1}^p \hat{\lambda}_k \nabla f_k(x^*) + \sum_{j \in \mathcal{J}(x^*)} \hat{\mu}_j \nabla g_j(x^*) = 0.$$

Definindo  $\hat{\mu}_j = 0 \quad \forall j \notin \mathcal{J}(x^*)$ , o teorema fica provado. ■

Uma consequência imediata deste teorema é apresentada no seguinte corolário.

**Corolário 2.2** *Se  $x^*$  é propriamente eficiente no sentido de Geoffrion e é cumprida a condição de qualificação KT em  $x^*$ , então as relações (2.26)-(2.29) são satisfeitas.*

O seguinte teorema relaciona as condições suficientes para pontos propriamente eficientes no sentido de Kuhn e Tucker e o problema escalar com pesos.

**Teorema 2.6** *Suponhamos que  $f_k, g_j : \mathbb{R}^n \rightarrow \mathbb{R}$  são funções convexas, continuamente diferenciáveis. Suponhamos que existem  $x^* \in \mathcal{X}$ ,  $\hat{\lambda} \in \mathbb{R}^p$  e  $\hat{\mu} \in \mathbb{R}^m$  satisfazendo (2.26)-(2.29). Então  $x^*$  é propriamente eficiente no sentido de Geoffrion.*

**Prova:** Consideremos a seguinte função escalar convexa  $f(x) = \sum_{k=1}^p \hat{\lambda}_k f_k(x)$ . Como  $x^*$  satisfaz (2.26)-(2.29), então é uma solução ótima global do problema

$$\min_{x \in \mathcal{X}} \sum_{k=1}^p \hat{\lambda}_k f_k(x).$$

Como  $\hat{\lambda}_k > 0 \quad \forall k = 1, \dots, p$ , o Teorema 2.2 nos diz que  $x^*$  é propriamente eficiente no sentido de Geoffrion. ■

O corolário deste teorema nos diz que para problemas convexos, ser propriamente eficiente no sentido de Kuhn e Tucker implica em ser propriamente eficiente no sentido de Geoffrion.

**Corolário 2.3** *Sejam  $f_k, g_j : \mathbb{R}^n \rightarrow \mathbb{R}$  funções convexas, continuamente diferenciáveis e suponhamos que  $x^*$  é propriamente eficiente no sentido de Kuhn e Tucker, então  $x^*$  é propriamente eficiente no sentido de Geoffrion.*

Após das definições dadas por Geoffrion e Kuhn-Tucker para pontos propriamente eficientes, surgiram outras definições dadas por Borwein (1977) e Benson (1979), as quais também estão relacionadas com a definição dada por Geoffrion. Para enunciá-las precisamos dos conceitos de cone tangente e casco cônico. Como o primeiro já foi enunciado (Definição 2.5) daremos a definição de casco cônico de um conjunto:

**Definição 2.7** *Seja  $\mathcal{K} \subset \mathbb{R}^p$ . O casco cônico de  $\mathcal{K}$  é definido como*

$$\text{cone}(\mathcal{K}) = \{\alpha y : \alpha \geq 0, y \in \mathcal{K}\} = \bigcup_{\alpha \geq 0} \alpha \mathcal{K}.$$

Para as seguintes definições chamaremos de  $\mathcal{Y}$  o espaço objetivo do problema multiobjetivo, isto é,  $\mathcal{Y} := \mathcal{F}(\mathcal{X})$ .

**Definição 2.8** *(Borwein (1977)) Uma solução  $x^* \in \mathcal{X}$  é propriamente eficiente no sentido de Borwein se*

$$\mathcal{T}_{\mathcal{Y} + \mathbb{R}_+^p}(f(x^*)) \cap (-\mathbb{R}_+^p) = \{0\}.$$

*(Benson (1979)) Uma solução  $x^* \in \mathcal{X}$  é propriamente eficiente no sentido de Benson se*

$$\text{cl}(\text{cone}(\mathcal{Y} + \mathbb{R}_+^p - f(x^*))) \cap (-\mathbb{R}_+^p) = \{0\}.$$

Podemos observar que estas definições não assumem que o ponto seja eficiente para ser ponto propriamente eficiente, mas pode-se provar que a eficiência é uma condição necessária para os pontos serem propriamente eficientes em ambos sentidos. As relações existentes são as seguintes: *se um ponto é propriamente eficiente no sentido de Benson então é propriamente eficiente no sentido de Borwein* e, sob hipóteses de convexidade ambas definições coincidem. Ademais *se um ponto é propriamente eficiente no sentido de Borwein, então o ponto é eficiente.*

Por outro lado, estas definições não usam a ordem das componentes da função vetorial, assim, nestas definições,  $\mathfrak{R}_+^p$  pode ser substituído por um cone convexo arbitrário  $\mathcal{C}$ . No entanto, a definição de Geoffrion utiliza a ordem nas componentes e, no caso que  $\mathcal{C} = \mathfrak{R}_+^p$ , as definições de Geoffrion e Benson coincidem. Para as provas das propriedades enunciadas anteriormente ver [15].

Para finalizar vamos destacar que as definições dadas por Benson e Borwein correspondem a condições dadas em termos de cones, enquanto a dada por Geoffrion não. Em consequência, a noção de eficiência própria de Geoffrion se mostra mais simples e mais clara.

Como neste trabalho estamos interessados nas soluções propriamente eficientes, baseados em todo o exposto, a definição dada por Geoffrion resulta a melhor caracterização, e portanto a definição que adotamos, a qual nos permite usar pesos estritamente positivos, que são mais úteis do ponto de vista algorítmico. No recente trabalho [8], Burachik e Rizvi estudam condições de otimalidade também relacionadas com pontos próprios no sentido de Geoffrion.

### 2.3.2 Conjuntos eficientes convexos

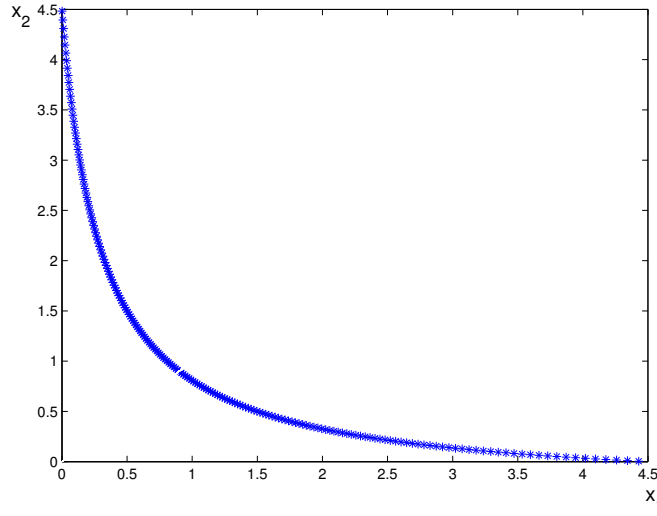
Em geral não é verdade que se as funções-objetivo e o domínio do problema multiobjetivo são convexos, o conjunto eficiente resulte convexo. Por exemplo se considerarmos o seguinte problema:

$$\begin{aligned} \min_{x \in \mathfrak{R}^2} \quad & (f_1(x), f_2(x)) \\ \text{s. a} \quad & x \geq 0. \end{aligned} \tag{2.33}$$

$$\text{onde } f_1(x) = x_1^2 + \frac{(x_2 - 9/2)^2}{4} \text{ e } f_2(x) = \frac{(x_1 - 9/2)^2}{4} + x_2^2,$$

o conjunto eficiente é apresentado no seguinte gráfico:





**Figura 2.4:** Conjunto de pontos eficientes  $\mathcal{X}^*$  correspondente ao problema dado em (2.33).

Neste exemplo, o conjunto eficiente é um conjunto não convexo. Mas quando as curvas de nível das funções objetivo são bolas no  $\mathbb{R}^n$ , o conjunto eficiente é um conjunto convexo. Isto é o que será discutido nesta seção.

Consideremos as funções escalares  $f_1, \dots, f_p$  com domínio no  $\mathbb{R}^n$  definidas como:

$$f_i(x) = \frac{\beta_i}{2} \|x - a_i\|^2 + \gamma_i, \quad \forall i = 1, \dots, p, \quad (2.34)$$

com  $a_i \in \mathbb{R}^n$ ,  $\beta_i \in \mathbb{R}_{++}$ ,  $\gamma_i \in \mathbb{R} \forall i = 1, \dots, p$ .

**Lema 2.1** *O fecho convexo determinado pelos pontos  $a_1, \dots, a_p$  é um conjunto eficiente do problema multiobjetivo  $\min_{x \in \mathbb{R}^n} (f_1, \dots, f_p)$ , em que as funções  $f_i$  são dadas em (2.34).*

**Prova:** Como as funções escalares  $f_1, \dots, f_p$  são convexas com hessiana definida positiva, múltiplos positivos delas também são convexas e a soma de funções convexas é uma função convexa, então para cada  $\alpha \in \mathbb{R}_+^p$  fixo com  $\alpha_1 + \dots + \alpha_p = 1$ , a função escalar  $\sum_{i=1}^p \alpha_i f_i(x)$  é convexa com hessiana definida positiva. Portanto o problema  $P(\alpha)$  :

$$\min_{x \in \mathbb{R}^n} \sum_{i=1}^p \alpha_i f_i(x) \quad (2.35)$$

tem solução única. Logo, pelo Corolário 2.1, para cada  $\alpha$  fixo, a solução do problema

$P(\alpha)$  é um ponto eficiente do problema multiobjetivo  $\min_{x \in \mathbb{R}^n} (f_1, \dots, f_p)$ .

Agora vamos procurar a solução ótima de  $P(\alpha)$ . Seja  $f(x) = \sum_{i=1}^p \alpha_i f_i(x)$ , das condições de otimalidade de primeira ordem para  $f$  temos:

$$\nabla f(x) = \sum_{i=1}^p \alpha_i \nabla f_i(x) = 0.$$

Então

$$\sum_{i=1}^p \alpha_i \beta_i (x - a_i) = 0.$$

Logo, como  $\sum_{j=1}^p \alpha_j \beta_j \neq 0$ , pois  $\sum_{i=1}^p \alpha_i = 1$ ,  $\alpha_i \geq 0$  e  $\beta_i > 0 \forall i = 1, \dots, p$ , temos

$$x = \frac{\sum_{i=1}^p \alpha_i \beta_i a_i}{\sum_{j=1}^p \alpha_j \beta_j} = \sum_{i=1}^p w_i a_i, \quad (2.36)$$

onde

$$w_i = \frac{\alpha_i \beta_i}{\sum_{j=1}^p \alpha_j \beta_j} \forall i = 1, \dots, p, \quad (2.37)$$

cumprindo  $w_1 + \dots + w_p = 1$  e  $w_i \geq 0 \forall i = 1, \dots, p$ . Portanto  $x$  é combinação convexa dos pontos  $a_1, \dots, a_p$ . Assim a solução do problema escalar  $P(\alpha)$  pertence ao fecho convexo determinado pelos pontos  $a_1, \dots, a_p$ . Suponhamos agora que temos  $x^*$  no fecho convexo determinado por  $a_1, \dots, a_p$ , então  $\exists w \in \mathbb{R}^p$ ,  $w \geq 0$ , tal que  $\sum_{i=1}^p w_i = 1$  e  $x^* = \sum_{i=1}^p w_i a_i$ . Queremos ver que existe  $\alpha \in \mathbb{R}_+^p$  tal que  $x^*$  é solução de  $P(\alpha)$ , isto é, devemos achar  $\alpha$  tal que

$$0 = \nabla f(x^*) = \sum_{i=1}^p \alpha_i \beta_i (x^* - a_i). \quad (2.38)$$

Escolhendo  $\alpha_i = \frac{w_i}{\beta_i \sum_{j=1}^p \frac{w_j}{\beta_j}} \forall i = 1, \dots, p$ , temos que  $\sum_{i=1}^p \alpha_i = 1$ ,  $\alpha_i \geq 0 \forall i = 1, \dots, p$  e  $x^*$  satisfaz (2.38) e assim,  $x^*$  corresponde à única solução de  $P(\alpha)$ .

Logo podemos dizer que todo ponto do fecho convexo determinado pelos pontos  $a_1, \dots, a_p$  é solução do problema  $P(\alpha)$  para algum  $\alpha \in \mathbb{R}^p$  com  $\alpha \geq 0$  e  $\alpha_1 + \dots + \alpha_p = 1$ . Portanto o conjunto solução do problema  $P(\alpha)$  está formado pelo fecho convexo determinado por  $a_1, \dots, a_p$ . Com isto fica provado o lema. ■

**Corolário 2.4** *O conjunto solução do problema  $P(\alpha)$ , dado em (2.35), para todo  $\alpha \geq 0$  tal que  $\alpha_1 + \dots + \alpha_p = 1$ , é o fecho convexo determinado pelos pontos  $a_1, \dots, a_p$ .*

**Corolário 2.5** *O interior do fecho convexo de  $a_1, \dots, a_p$ , é um conjunto propriamente eficiente do problema multiobjetivo  $\min_{x \in \mathbb{R}^n} (f_1, \dots, f_p)$ .*

**Prova:** Como  $f_i(x) \ \forall i = 1, \dots, p$ , são funções convexas e o domínio do problema multiobjetivo é convexo, pelos teoremas da seção anterior temos que os pontos propriamente eficientes do problema, são as soluções ótimas do problema escalar parametrizado em  $\alpha$ ,  $P(\alpha)$ , dado em (2.35) com  $\alpha_i > 0, \ \forall i = 1, \dots, p$ , e  $\sum_{i=1}^p \alpha_i = 1$ . Pelo Lema 2.1 as soluções de  $P(\alpha)$  pertencem ao fecho convexo determinado por  $a_1, \dots, a_p$ . Como  $\alpha > 0$  as soluções ótimas de  $P(\alpha)$  pertencem ao interior do fecho convexo. Assim os pontos propriamente eficientes do problema multiobjetivo pertencem ao interior do fecho convexo que os pontos  $a_1, \dots, a_p$  determinam.

Com estes resultados podemos enunciar o seguinte teorema:

**Teorema 2.7** *O conjunto solução do problema  $\min_{x \in \mathbb{R}^n} (f_1, \dots, f_p)$ , é o fecho convexo dos pontos  $a_1, \dots, a_p$ .*

**Prova:** A prova consta de duas partes:

1. Provar que o fecho convexo dos pontos  $a_1, \dots, a_p$  é solução do problema.
2. Não existe solução eficiente exterior ao politopo convexo determinado pelo fecho convexo .

A primeira parte fica provada pelo Lema 2.1. Assim só resta provar que pontos que pertencem ao exterior do politopo não podem ser solução do problema.

Suponhamos que  $z$  é um ponto exterior do politopo convexo determinado por  $a_1, \dots, a_p$ . Então existe um hiperplano  $\mathcal{H}$  que contém uma das faces do politopo e não contém o ponto  $z$ , isto é,  $z \notin \mathcal{H}$ . Seja  $w = \text{proj}_{\mathcal{H}} z$ , isto é, a projeção ortogonal de  $z$  sobre  $\mathcal{H}$ , então claramente  $z \neq w$ . Suponhamos que a face considerada contém  $r$  vértices do politopo. Sem perda de generalidade chamamos tais pontos de  $a_1, \dots, a_r$  e seja  $i$ , com  $1 \leq i \leq r$  arbitrário. Pela propriedade de  $w$  temos

$$\|w - a_i\| < \|z - a_i\| \Rightarrow f_i(w) < f_i(z),$$

como  $i$  é arbitrário, isto é satisfeito para todo  $i$  como  $1 \leq i \leq r$ , logo  $f_i(w) < f_i(z) \forall i = 1, \dots, r$ . Agora seja  $a_k$  outro vértice do politopo diferente de  $a_i \forall i = 1, \dots, r$ , e consideramos o triângulo de vértices  $a_k, w, z$ . Como o ângulo  $\widehat{vwz}$  é maior que  $\pi/2$ , temos que  $\|w - a_k\| < \|z - a_k\|$  então  $f_k(w) < f_k(z)$  para os restantes  $k$ . Portanto  $f_j(w) < f_j(z) \forall j = 1, \dots, p$ , assim  $z$  não é um ponto eficiente e o teorema fica provado. ■

Logo, retomando a discussão do começo desta seção, quando as funções objetivo do problema multiobjetivo estão dadas pela expressão (2.34), o conjunto eficiente corresponde a um conjunto convexo.

---

---

## Capítulo 3

# Restauração Inexata

Neste capítulo, incluído para completude do texto e comodidade do leitor, compilamos de Martínez [32] os principais resultados referentes ao método de Restauração Inexata (RI) para programação não linear, bem como o algoritmo de RI para programação em dois níveis proposto por Andreani et al. em [2].

### 3.1 Introdução

Consideremos o seguinte problema de Programação Não Linear (PNL):

$$\begin{array}{ll} \min & f(x) \\ \text{s. a} & C(x) = 0 \\ & x \in \Omega \end{array}$$

com  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  e  $C : \mathbb{R}^n \rightarrow \mathbb{R}^m$  funções de classe  $C^1(\mathbb{R}^n)$ ,  $\Omega$  conjunto convexo e compacto. Vamos chamar  $C'(x) \in \mathbb{R}^{m \times n}$  a matriz Jacobiana de  $C$  no ponto  $x$ .

Os métodos de Restauração Inexata (RI) são métodos iterativos, desenvolvidos para resolver problemas de PNL (ver [18], [31], [32]). Em cada iteração, a viabilidade e a otimalidade são analisadas em etapas diferentes.

Na primeira etapa do algoritmo é tratada a *viabilidade*, chamada *fase de Restauração*. Aqui, dado um ponto  $x^k \in \Omega$ , é procurado um ponto  $y^k \in \Omega$ , chamado de *ponto intermediário*, tal que a inviabilidade em  $y^k$  seja uma fração da inviabilidade em  $x^k$ . Na segunda etapa, *fase de Otimalidade*, é procurado um ponto  $z^{k,i} \in \pi_k \cap B_{k,i}$  onde  $\pi_k$  é um conjunto formado pelos pontos de  $\Omega$  tal que pertencem ao espaço que aproxima linearmente o espaço das restrições e,  $B_{k,i}$  representa uma região de confiança centrada em  $y^k$

de raio  $\delta_{k,i}$ , tal que  $f(z^{k,i}) \ll f(y^k)$ . Mais precisamente, pede-se que haja um decréscimo “suficiente” no valor de  $f$  em relação a  $y^k$ . No trabalho desenvolvido em [32] foi proposto que o  $z^{k,i}$  satisfaça:  $L(z^{k,i}, \lambda^k) \ll L(y^k, \lambda^k)$ , ou seja, procura-se um decréscimo no valor do Lagrangiano na região  $\pi_k \cap B_{k,i}$ , sendo o Lagrangiano associado ao problema de PNL definido como

$$L(x, \lambda) = f(x) + C(x)^T \lambda,$$

para todo  $x \in \Omega$  e  $\lambda \in \Re^m$ .

O ponto obtido nesta última fase, denominado de *ponto atual*, é comparado com  $x^k$  mediante uma função de mérito conveniente. Se houver um decréscimo suficiente desta última função, o ponto atual é aceito como o iterando  $x^{k+1}$ , senão, o raio da região de confiança é reduzido e é feita novamente a fase da otimização.

A grande vantagem deste método é que, em cada iteração, a inviabilidade e a otimalidade são tratadas em etapas diferentes, o que permite utilizar algoritmos apropriados que melhor se adaptem à resolução de cada subproblema.

## 3.2 Descrição do modelo algorítmico

### 3.2.1 Fase de restauração

Como já foi mencionado, dado um ponto  $x^k \in \Omega$ , o algoritmo procura um ponto intermediário  $y^k \in \Omega$  mais viável que  $x^k$ . Como medir essa maior inviabilidade? Para isto pede-se que  $y^k$  satisfaça duas condições:

1.  $\|C(y^k)\| \leq r\|C(x^k)\|$ ,  $r \in (0, 1)$ .
2.  $\|y^k - x^k\| \leq \beta\|C(x^k)\|$ ,  $\beta > 0$ .

Note que neste capítulo  $\|\cdot\|$  representa uma norma arbitrária.

Com a primeira destas condições, pede-se que a inviabilidade no ponto  $y^k$ , seja no máximo uma proporção da inviabilidade no ponto  $x^k$ . A segunda condição pede que a distância entre o ponto intermediário e o atual seja limitada pela medida de inviabilidade no ponto atual, de modo que, se o ponto  $x^k$  é viável, então deve ser escolhido  $y^k = x^k$ .

### 3.2.2 Aproximação linear

Após obter  $y^k$  na fase de restauração, é feita uma aproximação linear da região viável contendo  $y^k$ . Esta região é definida como:

$$\pi_k = \{z \in \Omega \mid C'(y^k)(z - y^k) = 0\}.$$

### 3.2.3 Fase de minimização

Nesta fase, procura-se um ponto  $z^{k,i} \in \pi_k \cap B_{k,i}$ , com

$$B_{k,i} = \{x \in \mathbb{R}^n \mid \|x - y^k\| \leq \delta_{k,i}\}$$

e  $\delta_{k,i}$ , tal que

$$L(z^{k,i}, \lambda^k) \ll L(y^k, \lambda^k).$$

Os raios  $\delta_{k,i}$  são reduzidos, até que é achado um ponto  $z^{k,i}$  tal que seu valor numa função de mérito seja suficientemente menor que o valor desta mesma função em  $x^k$ .

Antes da etapa da minimização é calculada a direção

$$d^{k,tan} = P_{\pi_k}(y^k - \eta \nabla L(x, \lambda_k)) - y^k$$

(gradiente projetado contínuo de  $L$  sobre  $\pi_k$ ),  $P_{\pi_k}(z)$  é a projeção ortogonal de  $z$  sobre  $\pi_k$ . Esta direção, que é de descida (será provado mais adiante), é utilizada para medir a otimalidade no ponto  $z$ . Por outro lado, se  $d^{k,tan} = 0$ , então é escolhido  $z^{k,i} = y^k$ , senão  $z^{k,i}$  é solução aproximada do seguinte subproblema:

$$\begin{aligned} \min \quad & L(x, \lambda^k) \\ \text{s. a} \quad & x \in \pi_k \cap B_{k,i}. \end{aligned} \tag{3.1}$$

A resolução de (3.1) pode ser feita utilizando algoritmos convenientes para resolver problemas de PNL com restrições lineares, desde que a bola  $B_{k,i}$  seja definida com a norma infinito.

### 3.2.4 Função de mérito

A função de mérito utilizada é

$$\psi(x, \lambda, \theta) = \theta L(x, \lambda) + (1 - \theta) \|C(x)\|,$$

com  $\theta \in (0, 1]$  como parâmetro de penalização. Esta função é um tipo de penalização exata. A escolha do parâmetro  $\theta$  em cada iteração depende de considerações práticas e teóricas.

Agora, o objetivo é determinar  $z^{k,i}$  e  $\lambda_{trial}^{k,i}$  tais que:

$$\psi(x^k, \lambda^k, \theta^k) > \psi(z^{k,i}, \lambda_{trial}^{k,i}, \theta^k)$$

ou, se chamamos

$$Ared_{k,i} = \theta_{k,i} [L(x^k, \lambda^k) - L(z^{k,i}, \lambda_{trial}^{k,i})] + (1 - \theta_{k,i}) [\|C(x^k)\| - \|C(z^{k,i})\|]$$

à *redução atual*, espera-se que

$$Ared_{k,i} > 0.$$

Por outro lado, para garantir a convergência, é preciso pedir mais que uma simples redução, é preciso uma redução *suficiente* da função de mérito que é definida pelo seguinte teste:

$$Ared_{k,i} \geq 0.1 Pred_{k,i} \tag{3.2}$$

onde  $Pred_{k,i}$  corresponde à redução *desejada* da função de mérito entre  $x^k$  e  $z^{k,i}$ :

$$\begin{aligned} Pred_{k,i}(\theta) &= \theta [L(x^k, \lambda^k) - L(z^{k,i}, \lambda^k) - C(y^k)^T (\lambda_{trial}^{k,i} - \lambda^k)] \\ &\quad + (1 - \theta) [\|C(x^k)\| - \|C(y^k)\|]. \end{aligned}$$

Dependendo do valor de  $\theta$ ,  $Pred_{k,i}(\theta)$  pode ser negativo, mas se  $\theta$  é suficientemente pequeno, então  $Pred_{k,i}$  está perto de  $[\|C(x^k)\| - \|C(y^k)\|]$ , e assim pode ser não negativo, portanto  $\exists \theta \in (0, 1]$  tal que

$$Pred_{k,i} \theta \geq \frac{1}{2} [\|C(x^k)\| - \|C(y^k)\|].$$



O maior  $\theta$  que cumpre esta condição é denotado por  $\theta_{k,i}$ . Por outro lado, se (3.2) é satisfeita, então define-se  $x^{k+1} = z^{k,i}$ . Senão, é reduzida a região de confiança e é feita a fase de minimização novamente.

### 3.2.5 Algoritmo 1

Sejam  $r \in [0, 1)$ ,  $\beta$ ,  $\eta$ ,  $\delta_{min}$ ,  $\tau_1$ ,  $\tau_2$  parâmetros positivos do algoritmo e, seja  $\{w_k\}$  tal que  $\sum_{i=1}^{\infty} w_k < \infty$ . Seja  $x^0 \in \Omega$  uma aproximação inicial,  $\theta_{-1} \in (0, 1)$  valor inicial do parâmetro de penalidade, e  $\delta_{1,0} > \delta_{min}$ . Faça  $k = 0$ .

**Passo 1** Calcular  $y^k \in \Omega$  tal que

$$\begin{aligned} \|C(y^k)\| &\leq r\|C(x^k)\| \\ \|y^k - x^k\| &\leq \beta\|C(x^k)\| \end{aligned}$$

Se não existir  $y^k$ , parar a execução do algoritmo.

Definir  $i = 0$ ,  $\delta_{k,0} \geq \delta_{min}$ ,  $\theta_{k,-1} = \min\{1, \min\{\theta_{-1}, \dots, \theta_{k-1}\} + w_k\}$

**Passo 2** Definir

$$\pi_k = \{z \in \Omega \mid C'(y^k)(z - y^k) = 0\}$$

Calcular

$$d_{tan}^k = P_k(y^k - \eta \nabla L(y^k, \lambda^k)) - y^k$$

$P_k(z)$  representa a projeção ortogonal de  $z$  em  $\pi_k$ .

Se  $\|d_{tan}^k\| = 0$  e  $y^k = x^k$  ou  $\|C(y^k)\| = 0$  parar a execução do algoritmo, retornando como  $y^k$  a “solução”.

**Passo 3** Se  $d_{tan}^k = 0$ , definir  $z^{k,i} = y^k$ ;

senão, achar  $z^{k,i} \in \pi_k$  tal que

$$L(z^{k,i}, \lambda^k) < L(y^k, \lambda^k)$$

$$s. a \quad \|z^{k,i} - y^k\| \leq \delta_{k,i}$$

**Passo 4** Se  $d_{tan}^k = 0$ , definir  $\lambda_{trial}^{k,i} = \lambda^k$ ,

senão, calcular  $\lambda_{trial}^{k,i} \in \Re^m$  tal que  $|\lambda_{trial}^{k,i}| \leq M$ .

**Passo 5** Definir  $\forall \theta \in [0, 1]$

$$Pred_{k,i}(\theta) = \theta[L(x^k, \lambda^k) - L(z^{k,i}, \lambda^k) - C(y^k)^T(\lambda_{trial}^{k,i} - \lambda^k)] + (1 - \theta)[\|C(x^k)\| - \|C(y^k)\|]$$

Calcular  $\theta_{k,i}$ , o máximo  $\theta \in [0, \theta_{k,i-1}]$  tal que

$$Pred_{k,i}(\theta) \geq \frac{1}{2}[\|C(x^k)\| - \|C(y^k)\|]$$

Definir

$$Pred_{k,i} = Pred_{k,i}(\theta_{k,i})$$

**Passo 6** Calcular

$$Ared_{k,i} = \theta_{k,i}[L(x^k, \lambda^k) - L(z^{k,i}, \lambda_{trial}^{k,i})] + (1 - \theta_{k,i})[\|C(x^k)\| - \|C(z^{k,i})\|]$$

Se

$$Ared_{k,i} \geq 0.1Pred_{k,i}$$

definir

$$x^{k+1} = z^{k,i}, \lambda^{k+1} = \lambda_{trial}^{k,i}, \theta_k = \theta_{k,i}, \delta_k = \delta_{k,i}, iacc(k) = i$$

$$Ared_k = Ared_{k,i}, Pred_k = Pred_{k,i}$$

e terminar a iteração  $k$ .

Senão, escolher  $\delta_{k,i+1} \in [0.1\delta_{k,i}, 0.9\delta_{k,i}]$ , definir  $i = i + 1$  e voltar ao Passo 4.

### 3.2.6 Observações

**Lema 3.1** A direção  $d^{k,tan}$  é direção de descida para a função Lagrangiana, no ponto  $y^k$ .

**Prova:** Como  $P_k(z)$  projeta ortogonalmente  $z$  em  $\pi_k$ , então

$$y^k - \eta \nabla L_x(y^k, \lambda^k) - P_k(y^k - \eta \nabla L_x(y^k, \lambda^k)) \perp \pi_k,$$

portanto

$$\|y^k - \eta \nabla L_x(y^k, \lambda^k) - P_k(y^k - \eta \nabla L_x(y^k, \lambda^k))\|^2 + \|y^k - P_k(y^k - \eta \nabla L_x(y^k, \lambda^k))\|^2 =$$

$$\begin{aligned} \|y^k - \eta \nabla L_x(y^k, \lambda^k) - P_k(y^k - \eta \nabla L_x(y^k, \lambda^k) - (y^k - P_k(y^k - \eta \nabla L_x(y^k, \lambda^k)))\|^2 = \\ \|\eta \nabla L_x(y^k, \lambda^k)\|^2. \end{aligned}$$

Por outro lado

$$\|y^k - \eta \nabla L_x(y^k, \lambda^k) - P_k(y^k - \eta \nabla L_x(y^k, \lambda^k))\| = \|d^{k,tan} + \eta \nabla L_x(y^k, \lambda^k)\|,$$

logo juntando as duas igualdades anteriores temos:

$$\|d^{k,tan} + \eta \nabla L_x(y^k, \lambda^k)\|^2 \leq \|\eta \nabla L_x(y^k, \lambda^k)\|^2. \quad (3.3)$$

Agora

$$\|d^{k,tan} + \eta \nabla L_x(y^k, \lambda^k)\|^2 = \|d^{k,tan}\|^2 + \|\eta \nabla L_x(y^k, \lambda^k)\|^2 + 2\eta(d^{k,tan})^T \nabla L_x(y^k, \lambda^k).$$

Utilizando (3.3) temos

$$\|d^{k,tan}\|^2 + \|\eta \nabla L_x(y^k, \lambda^k)\|^2 + 2\eta(d^{k,tan})^T \nabla L_x(y^k, \lambda^k) \leq \|\eta \nabla L_x(y^k, \lambda^k)\|^2,$$

assim

$$(d^{k,tan})^T \nabla L_x(y^k, \lambda_k) \leq -\frac{1}{2\eta} \|d^{k,tan}\|^2.$$

$\therefore d^{k,tan}$  é direção de descida.■

Observamos que  $\{\theta_k\}$  é uma sequência convergente. De fato, definindo

$$\theta_{k,small} = \min\{\theta_{-1}, \dots, \theta_k\}$$

e

$$\theta_{k,large} = \theta_{k,small} + w_k,$$

temos que  $\theta_{k+1} \leq \theta_{k,large}$  e  $\theta_k \geq \theta_{k,small} \forall k$ . Claramente  $\{\theta_{k,large}\}$  e  $\{\theta_{k,small}\}$  são convergentes para o mesmo limite, portanto  $\{\theta_k\}$  é convergente.

### 3.3 Hipóteses gerais e consequências

Para as provas da convergência do método é preciso fazer as seguintes suposições em relação ao problema:

**A.1** O conjunto  $\Omega$  é convexo e compacto.

**A.2** A matriz Jacobiana de  $C(x)$  existe e satisfaz a condição de Lipschitz

$$\exists L_1 > 0 \mid \|C'(y) - C'(x)\| \leq L_1 \|y - x\| \quad \forall x, y \in \Omega.$$

**A.3** O gradiente de  $f$  existe e satisfaz a condição de Lipschitz

$$\exists L_2 > 0 \mid \|\nabla f(y) - \nabla f(x)\| \leq L_2 \|y - x\| \quad \forall x, y \in \Omega.$$

Das condições anteriores segue-se que:

$$\|C(y) - C(x) - C'(x)(y - x)\| \leq \frac{L_1}{2} \|y - x\|^2 \quad \forall x, y \in \Omega, \quad (3.4)$$

$$|f(y) - f(x) - \nabla f(x)^T(y - x)| \leq \frac{L_2}{2} \|y - x\|^2 \quad \forall x, y \in \Omega, \quad (3.5)$$

$$|C_j(y) - C_j(x) - \nabla C_j^T(x)(y - x)| \leq \frac{L_1}{2} \|y - x\|^2 \quad \forall x, y \in \Omega. \quad (3.6)$$

**Prova:** Será feita a prova de (3.4), as provas de (3.5) e (3.6) são semelhantes.

Sabe-se que

$$\int_0^1 C'(x + t(y - x))(y - x) dt = \int_x^y C'(u) du = C(y) - C(x), \quad (3.7)$$

subtraindo a ambos os lados de (3.7)  $C'(x)(y - x)$  temos

$$\begin{aligned} C(y) - C(x) - C'(x)(y - x) &= \int_0^1 C'(x + t(y - x))(y - x) dt - \int_0^1 C'(x)(y - x) dt = \\ &= \int_0^1 [C'(x + t(y - x)) - C'(x)](y - x) dt \end{aligned}$$

então

$$\begin{aligned} \|C(y) - C(x) - C'(x)(y - x)\| &\leq \int_0^1 \|C'(x + t(y - x)) - C'(x)\| \|y - x\| dt \\ &\leq \int_0^1 L_1 \|t(y - x)\| \|y - x\| dt = \int_0^1 L_1 t \|y - x\|^2 dt = \frac{L_1}{2} \|y - x\|^2, \end{aligned}$$

portanto

$$\|C(y) - C(x) - C'(x)(y - x)\| \leq \frac{L_1}{2} \|y - x\|^2. \blacksquare$$

A seguir serão apresentados alguns resultados teóricos do método.

O seguinte resultado refere-se à deterioração que pode ser esperada na inviabilidade. Ele diz que podemos esperar um aumento limitado na inviabilidade em  $z^{k,i}$  (ponto obtido na fase de otimização) com relação à inviabilidade em  $y^k$ .

**Teorema 3.1** *Existe uma constante  $c_1 > 0$  (independente de  $k$ ), tal que quando  $y \in \Omega$  e  $z \in \pi_k$  temos que*

$$\|C(z)\| \leq \|C(y^k)\| + c_1 \|z - y^k\|^2.$$

**Prova:** Ver Lema 3.2 de [32].

O próximo teorema refere-se ao decréscimo que pode ser esperado de  $y^k$  a  $z^{k,i}$ . As constantes  $\tau_1$  e  $\tau_2$  presentes no teorema, correspondem a constantes fixas positivas presentes no *algoritmo 2.2* de [32].

**Teorema 3.2** *Existem  $c_2$  e  $c_3$  constantes estritamente positivas, tais que quando é definido  $y^k \in \Omega$  e é calculado  $z^{k,i}$  no passo da minimização do algoritmo de RI então*

$$L(z^{k,i}, \lambda^k) \leq L(y^k, \lambda^k) - \min\{\tau_2, c_2 \|d^{k,tan}\|^2, \tau_1 \delta_{k,i}, c_3 \|d^{k,tan}\| \delta_{k,i}\}.$$

**Prova:** Ver Lema 3.3 de [32].

A pergunta que surge do último teorema é se é possível achar sempre  $z^{k,i}$  tal que ele seja aceito na iteração, isto é tal que seja cumprido (3.2). A resposta é sim, escolhendo  $\delta_{k,i}$  suficientemente pequeno. O próximo teorema nos garante que o algoritmo está bem definido, assim é possível achar o  $\delta_{k,i}$  adequado.

**Teorema 3.3** *O algoritmo de RI está bem definido.*

**Prova:** Ver Teorema 4.1 de [32].

Com o seguinte teorema, temos um resultado ainda mais forte, ele diz que a sequência  $\{Ared_{k,iacc}\} \rightarrow 0$ .

**Teorema 3.4** *Suponhamos que o algoritmo de RI gera uma sequência infinita. Então*

$$\lim_{k \rightarrow \infty} Ared_k = 0.$$

**Prova:** Ver Lema 4.1 de [32].

Uma consequência deste teorema, é que a sequência das inviabilidade convergem para zero. Logo um ponto limite da sequência gerada pelo algoritmo é um ponto viável.

**Teorema 3.5** *Se o algoritmo de RI não para no passo 1  $\forall k = 0, 1, \dots$ , então*

$$\lim_{k \rightarrow \infty} \|C(x^k)\| = 0.$$

**Prova:**

$$\|C(x^k)\| = (1-r)\|C(x^k)\| + r\|C(x^k)\| \geq (1-r)\|C(x^k)\| + \|C(y^k)\|$$

então

$$\|C(x^k)\| \leq \frac{\|C(x^k)\| - \|C(y^k)\|}{1-r} \leq \frac{2}{1-r} Pred_{k,iacc} \leq \frac{2}{1-r} 0.1 Ared_{k,iacc(k)}$$

logo, utilizando o Teorema 3.4 temos  $\|C(x^k)\| \rightarrow 0$ . ■

Os resultados apresentados até aqui dizem que se o algoritmo de RI não para no primeiro passo, então é possível aproximar-se da região viável tanto quanto for desejado. Agora, o que acontece com a otimalidade? Isto é o que nos diz o próximo teorema:

**Teorema 3.6** *Seja  $\{x_k\}$  uma sequência infinita gerada pelo algoritmo de RI Então existe  $K_2$  um subconjunto infinito de  $\{0, 1, \dots\}$  tal que*

$$\lim_{k \in K_2} \|d^{k,tan}\| = 0.$$

**Prova:** A prova deste teorema pode ser encontrada em [32]. Ela é feita por absurdo, supõe que  $\exists k_0 > 0$  tal que a sequência  $\{\|d^{k,tan}\|\}$  está afastada do zero, isto é  $\|d^{k,tan}\| > 0 \forall k \geq k_0$ . Com isto é provado que a sequência dos parâmetros de penalidade  $\{\theta_k\}$  estão limitados inferiormente por um valor maior que zero. Finalmente, com isto é obtido que o resultado do Teorema 3.4 não é satisfeito, o que nos leva a um absurdo.■

Concluindo, se o algoritmo gera uma sequência que tem ponto limite, podemos garantir pelo Teorema 3.5 que ele é um ponto viável, e pelo Teorema 3.6, que ele é um ponto AGP. Do inglês *Approximate Gradient Projection*, esta condição significa que a sequência de aproximações do gradiente projetado converge a zero. Foi provado em [30], que esta propriedade é satisfeita pelos minimizadores locais de problemas de PNL com restrições, independentemente das condições de qualificação. Também foi provado que um ponto AGP é um ponto KKT sob condições bem abrangentes ([30], [32], [4], [5], [6], [20]).

### 3.4 Restauração Inexata para programação em dois níveis

O algoritmo de Restauração Inexata foi utilizado em [2] para resolver o seguinte tipo de problema:

$$\begin{aligned}
 \min_{x,y} \quad & F(x, y) \\
 \text{s. a} \quad & x \in X \\
 & y = \arg \min_y f(x, y), \\
 & \text{s. a} \quad h(x, y) = 0, \\
 & y \in Y,
 \end{aligned} \tag{3.8}$$

onde  $F : \mathbb{R}^{n+m} \rightarrow \mathbb{R}$ ,  $f : \mathbb{R}^{n+m} \rightarrow \mathbb{R}$ ,  $h : \mathbb{R}^{n+m} \rightarrow \mathbb{R}^p$  e os conjuntos  $X \subset \mathbb{R}^n$ ,  $Y \subset \mathbb{R}_+^m$ , são considerados como caixas. O problema de baixo, parametrizado em  $x$ , é chamado *problema de segundo nível* ou *problema de nível inferior* e, o problema que envolve minimizar em  $x$  e  $y$  é chamado *problema de primeiro nível* ou *problema de nível superior*. Notar que as restrições do problema de primeiro nível são definidas pela otimização de um segundo problema parametrizado em  $x$ . É por isto que se diz que os problemas de

programação em dois níveis têm uma estrutura de hierarquia.

Para a implementação do algoritmo assumimos  $Y = \Re_+^m$  e definimos:

$$C(x, y, \mu, \gamma) = \begin{bmatrix} \nabla_y f(x, y) + \nabla_y h(x, y)\mu - \gamma \\ h(x, y) \\ \gamma_1 y_1 \\ \vdots \\ \gamma_m y_m \end{bmatrix}$$

onde  $(\mu, \gamma) \in \Delta \subset \Re^p \times \Re^m$ . As condições  $KKT(x)$  do problema de segundo nível correspondem a

$$C(x, y, \mu, \gamma) = 0, \quad y \geq 0, \quad \gamma \geq 0.$$

O Lagrangiano para o problema, é definido como:

$$L(s, \lambda) = F(x, y) + C(s)^T \lambda$$

onde  $s = (x, y, \mu, \gamma)$ .

O método de *RI* para estes problemas gera a sequência  $\{s^k\} = \{(x^k, y^k, \mu^k, \gamma^k)\}$ , que é obtida com uma adaptação do Algoritmo 1, para problemas em dois níveis (ver [2]), o qual fica:

**Algoritmo 2:**

Sejam  $r \in [0, 1)$ ,  $\beta$ ,  $\eta$ ,  $\delta_{min}$ , parâmetros positivos do algoritmo e, seja  $\{w_k\}$  tal que  $\sum_{i=1}^{\infty} w_k < \infty$ . Seja  $x^0 \in \Omega$  uma aproximação inicial,  $\theta_{-1} \in (0, 1)$  valor inicial do parâmetro de penalidade, e  $\delta_{1,0} > \delta_{min}$ . Faça  $k = 0$ .

**Passo 1** Achar minimizador aproximado,  $y' \geq 0$  e os multiplicadores de Lagrange  $(\mu', \gamma') \in$



$\Delta$  do seguinte problema

$$\begin{aligned} \min_y \quad & f(x^k, y) \\ \text{s. a} \quad & h(x^k, y) = 0 \\ & y \geq 0 \end{aligned} \tag{3.9}$$

logo é definido como o ponto intermediário:  $v^k = (x^k, y', \mu', \gamma')$

Se não existir solução do problema (4.2), parar a execução do algoritmo.

Definir  $i = 0$ ,  $\delta_{k,0} \geq \delta_{\min}$ ,  $\theta_{k,-1} = \min\{1, \min\{\theta_{-1}, \dots, \theta_{k-1}\} + w_k\}$

**Passo 2** Definir

$$\begin{aligned} z &= (x, y, \lambda, \gamma), \text{ com } x \in X, y \in Y_+, \mu \in \Re^p, \gamma \in \Re_+^m \\ \Omega &= X \times Y_+ \times \Re^p \times \Re_+^m \\ \pi_k &= \{z \in \Omega \mid C'(v^k)(z - v^k) = 0\} \end{aligned}$$

Calcular

$$d_{tan}^k = P_k(v^k - \eta \nabla L(v^k, \lambda^k)) - v^k$$

$P_k(w)$  representa a projeção ortogonal de  $w$  em  $\pi_k$ .

Se  $\|d_{tan}^k\| = 0$  e  $v^k = s^k$  (ou  $\|C(v^k)\| = 0$ ) parar a execução do algoritmo, retornando como  $v^k$  a “solução” do problema *em dois níveis*.

**Passo 3** Se  $d_{tan}^k = 0$ , definir  $z^{k,i} = v^k$ ;

senão, achar  $z^{k,i} \in \pi_k$  tal que

$$L(z^{k,i}, \lambda^k) < L(v^k, \lambda^k)$$

$$\text{s. a } \|z^{k,i} - v^k\| \leq \delta_{k,i}$$

**Passo 4** Se  $d_{tan}^k = 0$ , definir  $\lambda_{trial}^{k,i} = \lambda^k$ ,

senão, calcular  $\lambda_{trial}^{k,i} \in \Re^m$  tal que  $|\lambda^{k,i}| \leq M$ .

**Passo 5** Definir  $\forall \theta \in [0, 1)$

$$Pred_{k,i}(\theta) = \theta[L(s^k, \lambda^k) - L(z^{k,i}, \lambda^k) - C(v^k)^T(\lambda_{trial}^{k,i} - \lambda^k)] + (1 - \theta)[\|C(s^k)\| - \|C(v^k)\|]$$

Calcular  $\theta_{k,i}$ , o máximo  $\theta \in [0, \theta_{k,i-1}]$  tal que

$$Pred_{k,i}(\theta) \geq \frac{1}{2}[\|C(s^k)\| - \|C(v^k)\|]$$

Definir

$$Pred_{k,i} = Pred_{k,i}(\theta_{k,i})$$

**Passo 6** Calcular

$$Ared_{k,i} = \theta_{k,i}[L(s^k, \lambda^k) - L(z^{k,i}, \lambda_{trial}^{k,i})] + (1 - \theta_{k,i})[\|C(s^k)\| - \|C(z^{k,i})\|]$$

Se

$$Ared_{k,i} \geq 0.1Pred_{k,i}$$

definir

$$s^{k+1} = z^{k,i}, \lambda^{k+1} = \lambda_{trial}^{k,i}, \theta_k = \theta_{k,i}, \delta_k = \delta_{k,i}, iacc(k) = i$$

$$Ared_k = Ared_{k,i}, Pred_k = Pred_{k,i}$$

e terminar a iteração  $k$ .

Senão, escolher  $\delta_{k,i+1} \in [0.1\delta_{k,i}, 0.9\delta_{k,i}]$ , definir  $i = i + 1$  e voltar ao Passo 4.

A análise de convergência deste algoritmo pode ser encontrada em [2].

---

---

## Capítulo 4

# Reformulação de um problema de PNL com restrições multiobjetivo

Neste capítulo fazemos a conexão entre os Capítulos 2 e 3 desta tese. Apresentamos a estrutura do problema de interesse, sua reformulação bem como os detalhes específicos dos três métodos utilizados para a sua resolução.

### 4.1 Introdução

O problema de interesse neste trabalho tem a seguinte estrutura:

$$\begin{array}{ll} \min_x & F(x) \\ \text{s. a} & x \in \mathcal{X}^* \end{array} \quad (4.1)$$

onde  $F : \Omega \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ ,  $\mathcal{X}^*$  corresponde a conjunto de pontos eficientes do problema multiobjetivo

$$\begin{array}{ll} \min_x & (f_1(x), \dots, f_p(x)) \\ \text{s. a} & h(x) = 0, \\ & x \geq 0, \end{array}$$

$f_i : \mathcal{A} \subseteq \Omega \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ ,  $\forall i = 1, \dots, p$ ,  $h$  função vetorial e todas funções de classe  $C^1$  em seus domínios.

Do capítulo anterior, seção 3.4, podemos observar que este problema não corresponde a um problema em dois níveis, pois neste problema o conjunto solução não é descrito por dois vetores e o problema de cima e de baixo dependem do mesmo vetor variável.

Por outro lado, o problema de baixo é um problema multiobjetivo e, como foi exposto

no Capítulo 2, com hipóteses adequadas, pode ser transformado em um problema escalar com pesos. Esta é a estratégia que usaremos para transformar o problema de nosso interesse em um problema em dois níveis.

O vetor solução do problema transformado, será dividido em dois, a primeira parte corresponderá ao vetor dos pesos que chamaremos de  $\alpha$  e, a segunda parte ao vetor de variáveis do problema original que denotaremos com  $x$ . O vetor  $x$  será a solução ótima do problema de segundo nível parametrizado no vetor dos pesos  $\alpha$ . Assim o problema transformado terá a estrutura de um problema em dois níveis.

Fazendo as mudanças correspondentes no problema (4.1) e dado um vetor de tolerâncias  $\epsilon \in \mathbb{R}_{++}^p$ , o problema transformado fica:

$$\begin{aligned}
 \min_{\alpha, x} \quad & F(x) \\
 \text{s. a} \quad & \alpha_i > \epsilon_i, \quad \forall i = 1, \dots, p \\
 & \sum_{i=1}^p \alpha_i = 1, \\
 & x = \arg \min_x \sum_{i=1}^p \alpha_i f_i(x) \\
 & \text{s. a} \quad h(x) = 0, \\
 & x \geq 0.
 \end{aligned} \tag{4.2}$$

Agora claramente, o problema (4.2) tem a estrutura de um problema em dois níveis, o vetor solução é dividido em dois  $(\alpha, x)$ , onde  $x$  corresponde à solução de um segundo problema parametrizado em  $\alpha$ .

Como as funções do segundo nível são convexas, pelos Teoremas 2.1 e 2.2 do Capítulo 2, temos garantido que a solução do problema escalar com pesos também será solução do problema multiobjetivo. Portanto, para achar a solução do problema (4.1) vamos resolver o problema (4.2) que chamaremos de *problema em dois níveis-multiobjetivo* (PDNM), fazendo referencia que a restrição corresponde ao conjunto de pontos eficientes de um problema multiobjetivo.

## 4.2 Métodos Implementados

Para a resolução do problema (4.2) foram implementados três métodos, os quais listamos abaixo. Em seguida detalhamos cada um deles:

**Método 1:** Restauração Inexata 1 (RI1)

**Método 2:** Restauração Inexata 2 (RI2).

**Método 3:** Reformulação KKT.

### 4.2.1 Método 1: Restauração Inexata 1

A primeira implementação refere-se ao método de Restauração Inexata para programação em dois níveis [2]. Chamando  $y$  ao ponto intermediário obtido na fase de *restauração* e  $w$  o vetor  $w = (\alpha, x, \mu, \gamma)$ , usaremos a seguinte notação:  $C(w) = 0$  às condições KKT do problema de segundo nível no ponto  $w$ ,  $\mathcal{C}(w) = 0$  o sistema de equações não lineares formado pelas restrição  $C(x) = 0$  junto com a restrição referente ao vetor dos alfas:  $\sum_{i=1}^p \alpha_i - 1 = 0$  e,  $\Upsilon = \mathbb{R}^p \times \Omega \times \Delta$  com  $(\mu, \gamma) \in \Delta \subset \mathbb{R}^m \times \mathbb{R}_+^n$ , na fase de *otimalidade* resolve-se o seguinte subproblema

$$\begin{aligned} \min_w \quad & L(w, \lambda) \\ \text{s. a} \quad & w \in \pi(y) \cap B(y) \\ & \sum_{i=1}^p \alpha_i = 1 \\ & x \geq 0, \alpha \geq \epsilon, \end{aligned}$$

onde  $L(w, \lambda) = F(x) + \mathcal{C}(w)^T \lambda$ ,  $\lambda \in \mathbb{R}^{2n+m+1}$ ,  $\pi(y) = \{z \in \Upsilon \mid C'(y)(z - y) = 0\}$ ,  $B(y)$  é uma região de confiança ao redor do ponto  $y$  e,  $\epsilon \in \mathbb{R}^p$  fixo.

### 4.2.2 Método 2: Restauração Inexata 2

O segundo método implementado é uma variante do primeiro. Usando a mesma notação que no método anterior, na fase de otimalidade, ao invés do Lagrangiano, minimiza-se a função objetivo do nível superior, isto é, resolve-se o seguinte subproblema:

$$\begin{aligned}
& \min_w F(w) \\
& \text{s. a } w \in \pi(y) \cap B(y) \\
& \sum_{i=1}^p \alpha_i = 1 \\
& x \geq 0, \alpha \geq \epsilon,
\end{aligned}$$

onde,  $w$  e  $\Upsilon$  estão definidos como no método anterior,  $\pi(y) = \{z \in \Upsilon \mid C''(y)(z-y) = 0\}$ ,  $B(y)$  é uma região de confiança ao redor do ponto  $y$  e,  $\epsilon \in \mathbb{R}_{++}^p$  fixo.

### 4.2.3 Método 3: Reformulação KKT

Esta terceira proposta corresponde ao método mais conhecido para programação em dois níveis. É o método usado desde varias décadas atrás até hoje. Ele consiste em minimizar a função do primeiro nível sujeita às condições KKT do segundo nível, ou seja, procura-se a solução do seguinte problema:

$$\begin{aligned}
& \min_{\alpha, x, \mu, \gamma} F(x) \\
& \text{s. a } x \in \Omega \\
& \sum_{i=1}^p \alpha_i = 1 \\
& \alpha \in \mathbb{R}^p, \gamma \in \mathbb{R}^n \\
& \alpha > \epsilon \\
& \sum_{i=1}^p \alpha_i \nabla f_i(x) + \nabla h(x) \mu - \gamma = 0 \\
& h(x) = 0 \\
& \gamma_1 x_1 = 0 \\
& \vdots \\
& \gamma_n x_n = 0 \\
& x_i \geq 0, \gamma_i \geq 0, \quad \forall i = 1, \dots, n,
\end{aligned}$$

**Observação:** Notar que as duas primeiras estratégias consideram a estrutura do problema de segundo nível, pois elas acham os pontos ótimos do problema (na fase de otimalidade) no conjunto de pontos propriamente eficientes, enquanto esta última proposta não. Mais ainda, não temos garantia que os pontos obtidos com o método 3 sejam solução do pro-

blema de segundo nível, isto é que sejam pontos propriamente eficientes do nível inferior.

### 4.3 Algoritmo RI-PDNM

Nesta seção expomos o algoritmo para o Método 2 utilizado em nossas implementações. Dos resultados obtidos utilizando os métodos 1 e 2, precisamos agregar um passo intermediário no algoritmo entre o passo 2 e 3. Fizemos esta mudança, pois nos testes notamos que para alguns dos problemas implementados, o ponto intermediário  $y^k$ , não melhorava o suficiente em relação ao ponto  $s^k$  obtido na iteração anterior, isto é, não satisfazia a condição exigida da redução da sua inviabilidade ser uma fração da inviabilidade  $s^k$  e logo, o algoritmo falhava no passo 5, pois o  $Pred$  não conseguia cumprir a condição requerida. Com esta mudança, foi possível que o ponto intermediário, em todos os experimentos realizados, satisfizesse a condição exigida. Chamamos este passo de **passo 2 intermediário**.

A seguir mostramos um dos problemas nos quais surgiu esta dificuldade e logo, com a mudança no algoritmo, conseguimos melhorar notavelmente os resultados. Este problema corresponde ao número 13 dos experimentos implementados que são apresentados no próximo capítulo.

$$\begin{array}{ll} \min_{x_1, x_2} & \left(x_1 - \frac{9}{2}\right)^2 + x_2^2 \\ \text{s. a} & x \in \mathcal{X}^* \end{array}$$

onde  $\mathcal{X}^*$  corresponde ao conjunto eficiente do seguinte problema multiobjetivo:

$$\begin{array}{ll} \min_{x_1, x_2} & \left( x_1^2 + \frac{\left(x_2 - \frac{9}{2}\right)^2}{4}, x_2^2 + \frac{\left(x_1 - \frac{9}{2}\right)^2}{4} \right) \\ \text{s. a} & -x_1 - x_2 + 2 \leq 0 \\ & x_1, x_2 \geq 0 \end{array}$$

A solução corresponde ao ponto:  $\left\{ \left( \frac{9}{2}, 0 \right) \right\}$ .

Na Tabela 4.1 mostramos os resultados obtidos com o algoritmo com e sem o passo

intermediário 2. Chamamos de **algoritmo modificado** o algoritmo com o passo intermediário e chamamos de **algoritmo original** sem esse passo. Os dados apresentados são: o ponto inicial para os pesos:  $\alpha^0$ , ponto inicial para as variáveis do problema original:  $x^0$ ,  $y_1^k$  corresponde ao ponto intermediário obtido na iteração  $k$  no passo 2 e  $y_2^k$  corresponde ao ponto intermediário obtido na iteração  $k$  no passo intermediário 2. Também mostramos os valores de  $\|C(s^k)\|$ ,  $\|C(y_1^k)\|$  e  $\|C(y_2^k)\|$  os quais correspondem ao valor da norma infinito da restrição  $C(z) = 0$  no ponto  $s^k$ , ponto atual,  $y_1^k$  e no ponto  $y_2^k$  respectivamente.

	$(\alpha^0, \mathbf{x}^0)$	$k$	$\ C(s^k)\ $	$\ C(y_1^k)\ $	$\ C(y_2^k)\ $
<b>Alg. original</b>	(0.5, 0.5, -10, 10)	5	9.1178E - 04	9.2168E - 04	—
	(0.73, 0.27, 10, 20)	6	0.0017	0.0018	—
	(0.23, 0.77, 10, 20)	4	5.4212E - 04	6.3760E - 04	—
<b>Alg. modificado</b>	(0.5, 0.5, -10, 10)	5	9.1178E - 04	9.2168E - 04	3.4655E - 06
	(0.73, 0.27, 10, 20)	6	0.0017	0.0018	2.7691E - 06
	(0.23, 0.77, 10, 20)	4	5.4212E - 04	6.3760E - 04	3.4637E - 06

**Tabela 4.1:** Tabela comparativa. Ponto intermediário

Para este exemplo particular, o algoritmo original falhou nos três casos mostrados e o algoritmo modificado convergiu na iteração seguinte de ter ingressado ao *passo 2 intermediário*, nos três casos. Os dados mostrados na Tabela 4.1 foram realizados com o Método 2, mas obtivemos resultados semelhantes com o Método 1.

A seguir é apresentado o algoritmo para o problema de programação em dois níveis-multiobjetivo usando o Método 2 de RI.

Para a descrição do algoritmo, definimos:

$$C(\alpha, x, \mu, \gamma) = \begin{bmatrix} \sum_{i=1}^p \alpha_i \nabla f_i(x) + \nabla h(x) \mu - \gamma \\ h(x) \\ \gamma_1 x_1 \\ \vdots \\ \gamma_n x_n \end{bmatrix},$$



onde  $(\mu, \gamma) \in \Delta \subset \Re^m \times \Re_+^n$ . As condições  $KKT(\alpha)$  para o problema de segundo nível, correspondem ao seguinte sistema não linear

$$C(\alpha, x, \mu, \gamma) = 0, \quad x \geq 0, \quad \gamma \geq 0.$$

**Algoritmo:**

Seja  $(\alpha^0, x^0, \mu^0, \gamma^0)$  uma aproximação inicial de  $\alpha$ ,  $x$ , e dos multiplicadores de Lagrange do segundo nível,  $\theta_{-1} \in (0, 1)$  valor inicial do parâmetro de penalidade,  $\delta_{1,0} > \delta_{min}$ ,  $\epsilon \in \Re_{++}^p$ ,  $tol_1$  tolerância para a viabilidade,  $tol_2$  a tolerância para a otimalidade,  $r \in (0, 1)$ ,  $kmax$  (quantidade máxima de iterações do algoritmo),  $\eta \in (0, 1)$ . Assumir que para  $k \in \{0, 1, 2, \dots\}$  foram calculados  $s^k \in \Omega \times \Delta$  e  $\theta_{k-1}, \theta_{k-2}, \dots, \theta_0$ . Os passos para obter  $s^{k+1} = (\alpha^{k+1}, x^{k+1}, \mu^{k+1}, \gamma^{k+1})$  são dados a seguir.

**Passo 1:** Cálculo de  $\theta$ ,

**Passo 2:** Achar o minimizador aproximado,  $x \in \Re^n$  e os multiplicadores de Lagrange  $(\mu, \gamma) \in \Delta$  do seguinte problema:

$$\begin{aligned} \min_x \quad & \sum_{i=1}^p \alpha_i^k f(x_i) \\ \text{s.a} \quad & h(x) = 0 \\ & x \geq 0 \end{aligned}$$

Se não existir a solução, parar a execução do algoritmo.

Definir  $y^k = (\alpha^k, x, \mu, \gamma)$ .

**Passo 2 intermediário:** Esta etapa intermediária é feita quando o ponto obtido no Passo 2 não conseguiu uma redução suficiente na viabilidade.

**Se**  $\|C(y^k)\| > r\|C(s^k)\|$

**se**  $\|C(s^k)\| < tol_1$

$$y^k = s^k$$

**senão** Resolver o seguinte problema

$$\min_{(x, \gamma, \mu)} \|C(\alpha^k, x, \mu, \gamma)\|^2$$

Definir  $y^k = (\alpha^k, x, \mu, \gamma)$ ,

$y^k$  é o ponto obtido na viabilidade.

**Passo 3:** Definir

$$\mathcal{C}' = \begin{bmatrix} d \\ C'(y^k) \end{bmatrix}$$

onde  $d \in \mathbb{R}^{p+2n+m}$  é um vector linha que tem as suas primeiras  $p$  componentes iguais a 1 e zero nas demais componentes, e corresponde ao gradiente transposto da restrição do primeiro nível:  $\sum_{i=1}^p \alpha_i - 1 = 0$ .

Chamando

$$\pi_k = \{y \in \Upsilon \mid \mathcal{C}'(y^k)(y - y^k) = 0\},$$

calcular

$$d_{tan}^k = P_k(y^k - \eta \nabla F(y^k)) - y^k,$$

$P_k(w)$  representa a projeção ortogonal de  $w$  em  $\pi_k$ , obtida resolvendo-se o seguinte subproblema

$$\begin{aligned} \min_w \quad & \frac{1}{2} \|w - (y^k - \eta \nabla F(y^k))\|^2 \\ \text{s. a} \quad & w \in \pi_k \end{aligned}$$

**Se**  $\|d_{tan}^k\| < tol_2$ , (testamos a otimalidade de  $y^k$ )

**se**  $\|y^k - s^k\| < tol_1$  ou  $\|\mathcal{C}(y^k)\| < tol_1$ , parar a execução do algoritmo, regressando como  $y^k$  a *solução* do problema *em dois níveis*.

**senão** definir  $i = 0$  e escolher  $\delta_{k,0} \geq \delta_{\min}$

**Passo 4:** **Se**  $\|d_{tan}^k\| > tol_2$

Achar  $v = (\alpha, x, \mu, \gamma)$  a solução do seguinte subproblema:

$$\begin{aligned}
 & \min_s \quad F(s) \\
 & \text{s. a} \quad s \in \pi_k, \\
 & \quad \sum_{j=1}^p \alpha_j = 1, \\
 & \quad \|y^k - s\| \leq \delta_k \\
 & \quad \epsilon_j < \alpha_j < 1 - \epsilon_j, \quad \forall j = 1, \dots, p \\
 & \quad x \geq 0.
 \end{aligned}$$

**Senão**

definir  $v = y^k$ .

**Passo 5:** Definir  $\forall \theta \in [0, 1]$

$$Pred_{k,i}(\theta) = \theta[(F(s^k) - F(v))] + (1 - \theta)[\|C(s^k)\| - \|C(y^k)\|]$$

Calcular  $\theta_{k,i}$ , o máximo  $\theta \in [0, \theta_{k,i-1}]$  tal que

$$Pred_{k,i}(\theta) \geq \frac{1}{2}[\|C(s^k)\| - \|C(y^k)\|]$$

Definir

$$Pred_{k,i} = Pred_{k,i}(\theta_{k,i})$$

**Passo 6:** Calcular

$$Ared_{k,i} = \theta_{k,i}[F(s^k) - F(v)] + (1 - \theta_{k,i})[\|C(s^k)\| - \|C(v)\|]$$

**se**

$$Ared_{k,i} \geq 0.1Pred_{k,i}$$

definir

$$s^{k+1} = v,$$

e terminar a iteração  $k$ , ir ao Passo 1.

**Senão**, escolher  $\delta_{k,i+1} \in [0.1\delta_{k,i}, 0.9\delta_{k,i}]$ , definir  $i = i + 1$  e voltar ao Passo 4.

---

---

## Capítulo 5

# Problemas testados

### 5.1 Introdução

Neste capítulo apresentamos os problemas testados, os quais dividimos em dois grupos. O primeiro grupo está formado pelos problemas nos quais a quantidade de funções objetivo do problema multiobjetivo (PM) é fixa. Neste grupo também a quantidade de variáveis do problema é fixa. Ao segundo grupo pertencem os problemas nos quais é possível variar a quantidade de funções objetivo do PM, ou seja, problemas nos quais é possível gerar uma família de problemas. Além disso, cada grupo é diferenciado por ter restrições no PM ou não, além da caixa.

Geramos os problemas 1 até 8, 10 até 13 e 20 para testar o desempenho dos métodos no caso que as funções objetivo do PM fossem convexas. Os demais problemas foram retirados de trabalhos já feitos ou, inspirados neles para testar o desempenho para o caso de não valer a convexidade.

### 5.2 Problemas testados

Como o problema de nosso interesse tem a seguinte estrutura:

$$\begin{array}{ll} \min_x & F(x) \\ \text{s. a} & x \in \mathcal{X}^* \end{array}$$

onde  $F : \Omega \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$  e  $\mathcal{X}^*$  corresponde ao conjunto de pontos eficientes de um problema multiobjetivo, listaremos os problemas mostrando primeiro o problema multiobjetivo e

logo a função objetivo  $F(x)$  para cada um dos problemas testados. Ademais usaremos a seguinte notação:  $x \in \Re^n$ , mas quando  $n = 1$  simplesmente diremos  $x \in \Re$ , o conjunto  $\mathcal{X}^*$  representa o conjunto das *soluções eficientes* ou *soluções Pareto-ótimas*, e o conjunto  $\Omega^*$  representa o conjunto solução do respectivo problema. Para alguns problemas, foram desenhadas as curvas de nível da função objetivo do problema de primeiro nível, junto com o conjunto de pontos propriamente eficientes.

### 5.2.1 Grupo 1 sem restrições

**Problema 1 :**

Problema multiobjetivo:

$$\min_{x \in \Re} (x^2, (x-1)^2)$$

Função objetivo:  $F(x) = x$ ,

$$\mathcal{X}^* = [0, 1],$$

$$\Omega^* = \{0\}.$$

**Problema 2 :**

Problema multiobjetivo:

$$\min_{x \in \Re} (x^2, (x-1)^2, (x-2)^2)$$

Função objetivo:  $F(x) = x$ ,

$$\mathcal{X}^* = [0, 2],$$

$$\Omega^* = \{0\}.$$

**Problema 3 :**

Problema multiobjetivo:

$$\min_{x \in \Re^2} (x_1^2 + x_2^2, (x_1-1)^2 + (x_2-1)^2)$$

Função objetivo:  $F(x) = x_1 + x_2$ ,

$$\mathcal{X}^* = \{(x_1, x_2) \mid 0 \leq x_1 \leq 1, x_2 = x_1\},$$

$$\Omega^* = \{(0, 0)\}.$$

**Problema 4 :**

Problema multiobjetivo:

$$\min_{x \in \mathbb{R}^2} (x_1^2 + x_2^2, (x_1 - 1)^2 + (x_2 - 1)^2, (x_1 - 2)^2 + (x_2 - 1)^2, (x_1 - 1)^2 + (x_2 + 2)^2)$$

Função objetivo:  $F(x) = x_1 + x_2$ ,

$\mathcal{X}^* =$  quadrilátero de vértices  $(0, 0)$ ,  $(1, 1)$ ,  $(2, 1)$ ,  $(1, -2)$ ,

$$\Omega^* = \{(1, -2)\}.$$

**Problema 5 :**

Problema multiobjetivo:

$$\min_{x \in \mathbb{R}^2} (x_1^2 + x_2^2, (x_1 - 1)^2 + (x_2 + 1)^2)$$

Função objetivo:  $F(x) = x_1 + x_2$ ,

$\mathcal{X}^* = \{(x_1, x_2) \mid 0 \leq x_1 \leq 1, x_2 = -x_1\}$ ,

$\Omega^* = \{(x_1, x_2) \mid 0 \leq x_1 \leq 1, x_2 = -x_1\}$ .

**Problema 6 :**

Problema multiobjetivo:

$$\min_{x \in \mathbb{R}^2} (x_1^2 + x_2^2, (x_1 - 1)^2 + (x_2 - 1)^2, (x_1 - 2)^2 + (x_2 - 1)^2, (x_1 - 1)^2 + (x_2 + 1)^2)$$

Função objetivo:  $F(x) = x_1 + x_2$ ,

$\mathcal{X}^* =$  quadrilátero de vértices  $(0, 0)$ ,  $(1, 1)$ ,  $(2, 1)$ ,  $(1, -1)$ ,

$\Omega^* = \{(x_1, x_2) \mid 0 \leq x_1 \leq 1, x_2 = -x_1\}$ .

**Problema 7 :**

Problema multiobjetivo:

$$\min_{x \in \mathbb{R}^3} (\{(x_1 - 1)^2 + x_2^2 + x_3^2\}, \{x_1^2 + (x_2 - 1)^2 + x_3^2\}, \{(x_1 + 1)^2 + x_2^2 + x_3^2\}, \\ \{x_1^2 + (x_2 + 1)^2 + x_3^2\}, \{x_1^2 + x_2^2 + (x_3 - 1)^2\})$$

Função objetivo:  $F(x) = (x_3 - 1)^2$ ,

$\mathcal{X}^*$  = pirâmide de vértices  $(1, 0, 0)$ ,  $(0, 1, 0)$ ,  $(-1, 0, 0)$ ,  $(0, -1, 0)$ ,  $(0, 0, 1)$ ,

$\Omega^* = \{(0, 0, 1)\}$ .

### Problema 8 :

Problema multiobjetivo:

$$\min_{x \in \mathbb{R}^3} (\{(x_1 - 1)^2 + x_2^2 + x_3^2\}, \{x_1^2 + (x_2 - 1)^2 + x_3^2\}, \{(x_1 + 1)^2 + x_2^2 + x_3^2\}, \\ \{x_1^2 + (x_2 + 1)^2 + x_3^2\}, \{x_1^2 + x_2^2 + (x_3 - 1)^2\})$$

Função objetivo:  $F(x) = x_3$ ,

$\mathcal{X}^*$  = pirâmide de vértices  $(1, 0, 0)$ ,  $(0, 1, 0)$ ,  $(-1, 0, 0)$ ,  $(0, -1, 0)$ ,  $(0, 0, 1)$ ,

$\Omega^*$  = face da pirâmide na qual  $x_3 = 0$ .

### 5.2.2 Grupo 1 com restrições

No seguinte problema testado, as funções objetivo não são convexas nem diferenciáveis. Ele foi extraído de Zitzler E., Deb K., Thiele L. ([35]) e corresponde ao problema ZDT1:

#### Problema 9 ZDT1:

Problema multiobjetivo:

$$\min_{x \in \mathbb{R}^n} (f_1(x), f_2(x)) \\ s. a \quad x_i \in [0, 1] \quad \forall i = 1, \dots, n$$

onde

$$\begin{aligned} f_1(x_1) &= x_1 \\ f_2(x) &= g(x_2, \dots, x_n)h(f_1, g) \\ g(x_2, \dots, x_n) &= 1 + 9 \sum_{i=2}^n x_i / (n-1) \\ h(f_1, g) &= 1 - \sqrt{\frac{f_1}{g}} \\ x_i &\in [0, 1] \forall i = 1, \dots, n \end{aligned}$$

**Função objetivo para  $n = 1$ :**  $F(x) = (x - 1)^2$ ,

$$\mathcal{X}^* : \{x \in \mathbb{R} \mid x \in [0, 1]\},$$

$$\Omega^* : \{1\}.$$

**Função objetivo para  $n = 2$ :**  $F(x_1, x_2) = (x_1 - 1)^2 + x_2^2$ ,

$$\mathcal{X}^* : \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 \in [0, 1], x_2 = 0\},$$

$$\Omega^* : \{(1, 0)\}.$$

Nos próximos problemas, o PM possui restrições de desigualdade em sua formulação original. Com o acréscimo de folgas não negativas, chega-se ao formato (4.1), que é resolvido por meio do problema (4.2).

**Problema 10 :**

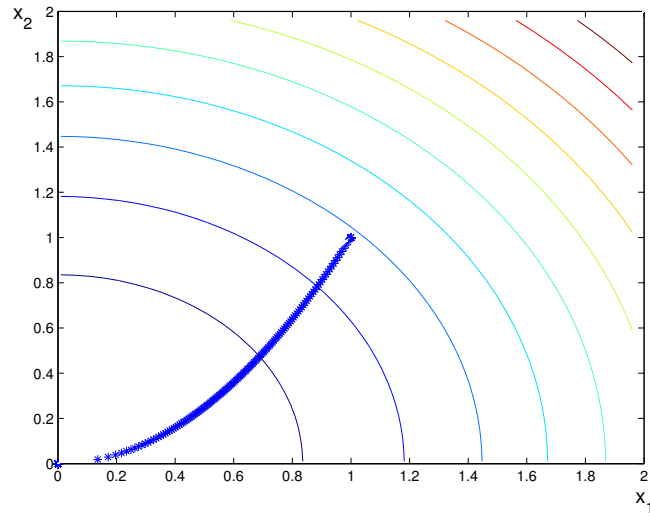
Problema multiobjetivo:

$$\begin{aligned} \min_{x \in \mathbb{R}^2} \quad & (-x_1, x_1 + x_2^2) \\ \text{s. a} \quad & x_1^2 - x_2 \leq 0 \\ & x_1 + 2x_2 \leq 3 \\ & x_1, x_2 \geq 0 \end{aligned}$$

Função objetivo:  $F(x) = x_1^2 + x_2^2$ ,

$$\mathcal{X}^* = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 \in [0, 1], x_2 = x_1^2\} \text{ (cf.Figura 5.1).}$$





**Figura 5.1:** Conjunto  $\mathcal{X}^*$  do problema 10.

$$\Omega^* = \{(0, 0)\}.$$

**Problema 11 :**

Problema multiobjetivo:

$$\begin{aligned} \min_{x \in \mathbb{R}^2} \quad & (-x_1, x_1 + x_2^2) \\ \text{s. a} \quad & x_1^2 - x_2 \leq 0 \\ & x_1 + 2x_2 \leq 3 \\ & x_1, x_2 \geq 0 \end{aligned}$$

Função objetivo:  $F(x) = (x_1 - 1)^2 + (x_2 - 1)^2$ ,

$$\mathcal{X}^* = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 \in [0, 1], x_2 = x_1^2\}$$

$$\Omega^* = \{(1, 1)\}.$$

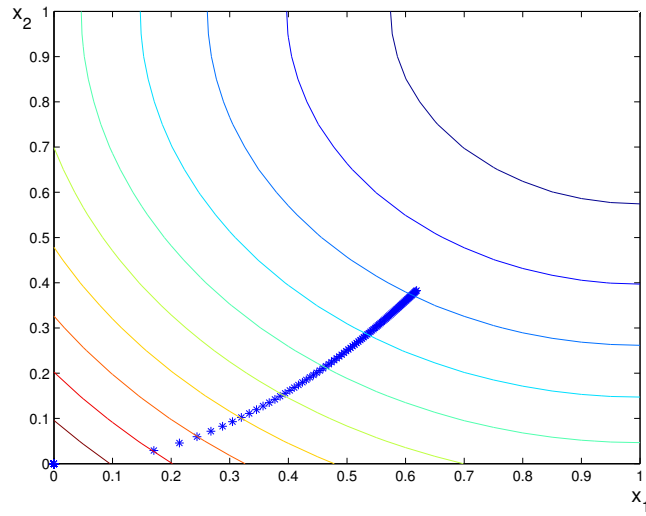
**Problema 12 :**

Problema multiobjetivo:

$$\begin{aligned} \min_{x \in \mathbb{R}^2} \quad & (x_1, -x_1 + x_2^2) \\ \text{s. a} \quad & x_1^2 - x_2 \leq 0 \\ & -x_1 + 2x_2 \leq 3 \\ & x_1, x_2 \geq 0 \end{aligned}$$

Função objetivo:  $F(x) = (x_1 - 1)^2 + (x_2 - 1)^2$ ,

$$\mathcal{X}^* = \left\{ (x_1, x_2) \in \mathbb{R}^2 \mid x_1 \in \left[0, \frac{2^{1/3}}{2}\right], x_2 = x_1^2 \right\} \text{ (cf. Figura 5.2).}$$



**Figura 5.2:** Conjunto  $\mathcal{X}^*$  do problema 12.

$$\Omega^* \approx \{(0.63, 0.39)\}.$$

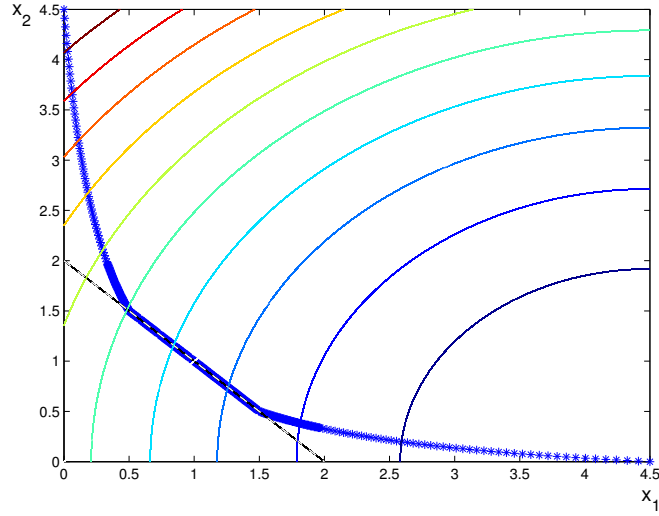
**Problema 13 :**

Problema multiobjetivo:

$$\begin{aligned} \min_{x \in \mathbb{R}^2} & \left( x_1^2 + \frac{\left(x_2 - \frac{9}{2}\right)^2}{4}, x_2^2 + \frac{\left(x_1 - \frac{9}{2}\right)^2}{4} \right) \\ \text{s. a} & \quad -x_1 - x_2 + 2 \leq 0 \\ & \quad x_1, x_2 \geq 0 \end{aligned}$$

Função objetivo:  $F(x) = \left(x_1 - \frac{9}{2}\right)^2 + x_2^2$ ,

$\mathcal{X}^*$  : O conjunto de pontos eficientes corresponde à curva desenhada com estrelas em azul (cf.Figura 5.3).



**Figura 5.3:** Conjunto  $\mathcal{X}^*$  do problema 13.

$$\Omega^* = \left\{ \left( \frac{9}{2}, 0 \right) \right\}.$$

Nos problemas exibidos até aqui, com exceção do problema 9, o problema multiobjetivo satisfaz as hipóteses dos Teoremas 2.1 e 2.2, assim o problema multiobjetivo tem funções coordenadas convexas e, a região viável é convexa. Logo, a solução do problema com pesos será o conjunto de pontos propriamente eficientes do segundo nível.

Como foi mencionado, o problema 9 não satisfaz as hipóteses dos Teoremas 2.1 e 2.2, a função objetivo  $f_2$  do PM não é convexa nem diferenciável, mas vamos ver nos resultados dos experimentos que mesmo assim, com os Métodos 1 e 2 conseguimos obter a solução do problema.

### 5.2.3 Grupo 2 sem restrições

Os problemas a seguir correspondem às famílias de problemas testadas.

As primeiras três foram inspiradas na família WFG de Huband S. et. al ([21], [22]). Foram feitas escolhas particulares dos parâmetros para trabalhar com funções diferenciáveis. Os problemas 17, 18 e 19 correspondem às famílias DTLZ1, DTLZ2 e DTLZ3 de Deb. K. et. al ([10]). Vamos destacar que as funções objetivo do PM de cada uma destas famílias não são convexas.

**Problema 14 :**

Problema multiobjetivo:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & (f_1(x), \dots, f_p(x)) \\ \text{s. a} \quad & x_i \in [0, 1] \quad \forall i = 1, \dots, n \end{aligned}$$

onde

$$n = p,$$

$$\begin{aligned} f_1(x) &= \frac{x_n}{2n} + 2 \prod_{i=1}^{p-1} \left( 1 - \cos \left( \frac{x_i \pi}{4i} \right) \right) \\ f_m(x) &= \frac{x_n}{2n} + 2m \prod_{i=1}^{p-m} \left( 1 - \cos \left( \frac{x_i \pi}{4i} \right) \right) \left( 1 - \sin \left( \frac{x_{p-m+1} \pi}{4(p-m+1)} \right) \right), \quad m = 2, \dots, p-1 \\ f_p(x) &= \frac{x_n}{2n} + 2p \left( 1 - \frac{x_1}{2} - \frac{\cos(5\pi x_1 + \pi/2)}{10\pi} \right) \end{aligned}$$

$$\text{Função objetivo: } F(x) = \sum_{i=1}^n x_i^2, \quad \Omega^* = \{\mathbf{0}\}.$$

**Problema 15 :**

Problema multiobjetivo:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & (f_1(x), \dots, f_p(x)) \\ \text{s. a} \quad & x_i \in [0, 1] \quad \forall i = 1, \dots, n \end{aligned}$$

onde

$$n = p,$$

$$\begin{aligned} f_1(x) &= \frac{x_n}{2n} + 2 \prod_{i=1}^{p-1} \left( 1 - \cos \left( \frac{x_i \pi}{4i} \right) \right) \\ f_m(x) &= \frac{x_n}{2n} + 2m \prod_{i=1}^{p-m} \left( 1 - \cos \left( \frac{x_i \pi}{4i} \right) \right) \left( 1 - \sin \left( \frac{x_{p-m+1} \pi}{4(p-m+1)} \right) \right), \quad m = 2, \dots, p-1 \\ f_p(x) &= \frac{x_n}{2n} + 2p \left( 1 - \frac{x_1}{2} \cos^2 \left( \frac{5\pi x_1}{2} \right) \right) \end{aligned}$$

$$\text{Função objetivo: } F(x) = \sum_{i=1}^n x_i^2, \quad \Omega^* = \{\mathbf{0}\}.$$

**Problema 16 :**

Problema multiobjetivo:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & (f_1(x), \dots, f_p(x)) \\ \text{s. a} \quad & x_i \in [0, 1] \quad \forall i = 1, \dots, n \end{aligned}$$

onde

$$\begin{aligned} n &= p, \\ f_1(x) &= \frac{x_n}{2n} + 2 \prod_{i=1}^{p-1} \sin\left(\frac{x_i \pi}{4i}\right) \\ f_m(x) &= \frac{x_n}{2n} + 2m \prod_{i=1}^{p-m} \sin\left(\frac{x_i \pi}{4i}\right) \cos\left(\frac{x_{p-m+1} \pi}{4(p-m+1)}\right), \quad m = 2, \dots, p-1 \\ f_p(x) &= \frac{x_n}{2n} + 2p \cos\left(\frac{x_1 \pi}{4}\right) \end{aligned}$$

Função objetivo:  $F(x) = \sum_{i=1}^n x_i^2$ ,  $\Omega^* = \{\mathbf{0}\}$ .

**Problema 17 DTLZ1::**

Problema multiobjetivo:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & (f_1(x), \dots, f_p(x)) \\ \text{s. a} \quad & x_i \in [0, 1] \quad \forall i = 1, \dots, n \end{aligned}$$

onde

$$\begin{aligned} n &\geq p, \\ f_1(x) &= (1 + g(x)) 0.5 \prod_{i=1}^{p-1} x_i \\ f_m(x) &= (1 + g(x)) 0.5 \prod_{i=1}^{p-m} x_i (1 - x_{p-m+1}), \quad m = 2, \dots, p-1 \\ f_p(x) &= (1 + g(x)) 0.5 (1 - x_1) \\ g(x) &= 100 \left[ p + \sum_{i=p}^n (x_i - 0.5)^2 - \cos(20\pi(x_i - 0.5)) \right] \end{aligned}$$

Função objetivo:  $F(x) = \sum_{i=1}^n x_i^2$ ,  $\Omega^* = \{\mathbf{0}\}$ .

**Problema 18** DTLZ2::

Problema multiobjetivo:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & (f_1(x), \dots, f_p(x)) \\ \text{s. a} \quad & x_i \in [0, 1] \quad \forall i = 1, \dots, n \end{aligned}$$

onde

$$\begin{aligned} n &\geq p, \\ f_1(x) &= (1 + g(x)) \prod_{i=1}^{p-1} \cos(x_i \pi / 2) \\ f_m(x) &= (1 + g(x)) \prod_{i=1}^{p-m} \cos(x_i \pi / 2) \sin(x_{p-m+1} \pi / 2), \quad m = 2, \dots, p-1 \\ f_p(x) &= (1 + g(x)) \sin(x_1 \pi / 2) \\ g(x) &= \sum_{i=p}^n (x_i - 0.5)^2 \end{aligned}$$

Função objetivo:  $F(x) = \sum_{i=1}^n x_i^2$ ,  $\Omega^* = \{\mathbf{0}\}$ .

**Problema 19** DTLZ3::

Problema multiobjetivo:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & (f_1(x), \dots, f_p(x)) \\ \text{s. a} \quad & x_i \in [0, 1] \quad \forall i = 1, \dots, n \end{aligned}$$

onde

$$\begin{aligned} n &\geq p, \\ f_1(x) &= (1 + g(x)) \prod_{i=1}^{p-1} \cos(x_i \pi / 2) \\ f_m(x) &= (1 + g(x)) \prod_{i=1}^{p-m} \cos(x_i \pi / 2) \sin(x_{p-m+1} \pi / 2), \quad m = 2, \dots, p-1 \\ f_p(x) &= (1 + g(x)) \sin(x_1 \pi / 2) \\ g(x) &= 100 \left[ p + \sum_{i=p}^n (x_i - 0.5)^2 - \cos(20\pi(x_i - 0.5)) \right] \end{aligned}$$

Função objetivo:  $F(x) = \sum_{i=1}^n x_i^2$ ,  $\Omega^* = \{\mathbf{0}\}$ .

### 5.2.4 Grupo 2 com restrições

As seguintes famílias, apresentam restrições no nível inferior. A primeira delas é um problema bem geométrico. O conjunto eficiente corresponde à interseção de um polígono regular de  $p$  lados, inscrito em um círculo com centro em  $(1, 1)$  e raio 2, e semiplanos.

**Problema 20 :**

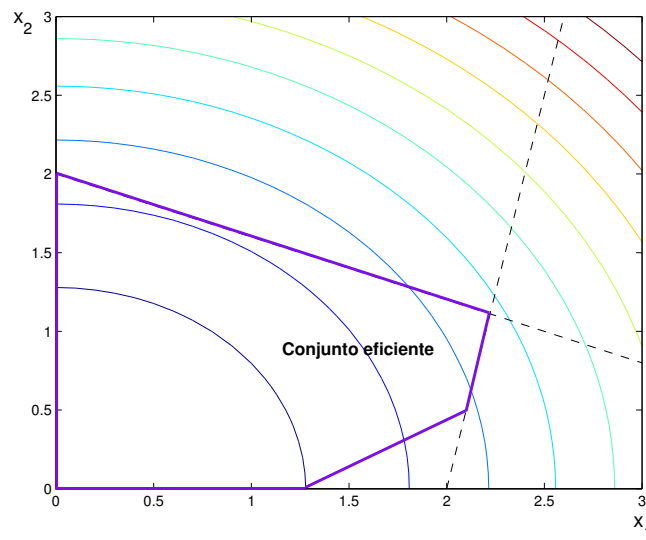
Problema multiobjetivo:

$$\begin{aligned} \min_{x_1, x_2} \quad & (f_1(x), \dots, f_p(x)) \\ \text{s. a} \quad & x_2 + \frac{2}{5}x_1 \leq 2 \\ & 5x_1 - x_2 \leq 10 \\ & x_1, x_2 \geq 0 \end{aligned}$$

onde

$$\begin{aligned} f_i(x_1, x_2) &= (x_1 - a_i)^2 + (x_2 - b_i)^2, \quad \forall i = 1, \dots, p \\ a_i &= 2 \cos(\theta_i) + 1 \quad \forall i = 1, \dots, p \\ b_i &= 2 \sin(\theta_i) + 1 \quad \forall i = 1, \dots, p \\ \theta_i &= \frac{2\pi}{p}(i - 1) \end{aligned}$$

Função objetivo:  $F(x) = x_1^2 + x_2^2$ ,  $\Omega^* = \{(0, 0)\}$ .



**Figura 5.4:** Ilustração de  $\mathcal{X}^*$  para o problema 20, com  $p = 3$ .

As duas famílias que seguem são uma variante do problema 14, o nível inferior apresenta uma restrição linear.

**Problema 21 :**

Problema multiobjetivo:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & (f_1(x), \dots, f_p(x)) \\ \text{s. a} \quad & \sum_{i=1}^n x_i \leq 1 \\ & x_i \in [0, 1] \quad \forall i = 1, \dots, n \end{aligned}$$

onde

$$n = p,$$

$$f_1(x) = \frac{x_n}{2n} + 2 \prod_{i=1}^{p-1} \left( 1 - \cos \left( \frac{x_i \pi}{4i} \right) \right)$$

$$f_m(x) = \frac{x_n}{2n} + 2m \prod_{i=1}^{p-m} \left( 1 - \cos \left( \frac{x_i \pi}{4i} \right) \right) \left( 1 - \sin \left( \frac{x_{p-m+1} \pi}{4(p-m+1)} \right) \right), \quad m = 2, \dots, p-1$$

$$f_p = \frac{x_n}{2n} + 2p \left( 1 - \frac{x_1}{2} - \frac{\cos(5\pi x_1 + \pi/2)}{10\pi} \right)$$

Função objetivo:  $F(x) = \sum_{i=1}^n x_i^2$ ,  $\Omega^* = \{0\}$ .

**Problema 22 :**

Problema multiobjetivo:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & (f_1(x), \dots, f_p(x)) \\ \text{s. a} \quad & \sum_{i=1}^n x_i \geq 1 \\ & x_i \in [0, 1] \quad \forall i = 1, \dots, n \end{aligned}$$



onde

$$n = p,$$

$$f_1(x) = \frac{x_n}{2n} + 2 \prod_{i=1}^{p-1} \left( 1 - \cos \left( \frac{x_i \pi}{4i} \right) \right)$$

$$f_m(x) = \frac{x_n}{2n} + 2m \prod_{i=1}^{p-m} \left( 1 - \cos \left( \frac{x_i \pi}{4i} \right) \right) \left( 1 - \sin \left( \frac{x_{p-m+1} \pi}{4(p-m+1)} \right) \right), \quad m = 2, \dots, p-1$$

$$f_p = \frac{x_n}{2n} + 2p \left( 1 - \frac{x_1}{2} - \frac{\cos(5\pi x_1 + \pi/2)}{10\pi} \right)$$

Função objetivo:  $F(x) = \sum_{i=1}^n x_i^2$ , para  $n = 2$ ,  $\Omega^* = \{(1, 0)\}$ .

As famílias a seguir, correspondem às famílias DTLZ1, DTLZ2 e DTLZ3 (problemas 17, 18 e 19 respectivamente), agregando uma restrição linear de desigualdade no nível inferior.

**Problema 23** DTLZ1-Res:

Problema multiobjetivo:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & (f_1(x), \dots, f_p(x)) \\ \text{s. a} \quad & \sum_{i=1}^n x_i \leq 1 \\ & x_i \in [0, 1] \quad \forall i = 1, \dots, n \end{aligned}$$

onde

$$n \geq p,$$

$$f_1(x) = (1 + g(x)) 0.5 \prod_{i=1}^{p-1} x_i$$

$$f_m(x) = (1 + g(x)) 0.5 \prod_{i=1}^{p-m} x_i (1 - x_{p-m+1}), \quad m = 2, \dots, p-1$$

$$f_p(x) = (1 + g(x)) 0.5 (1 - x_1)$$

$$g(x) = 100 \left[ p + \sum_{i=p}^n (x_i - 0.5)^2 - \cos(20\pi(x_i - 0.5)) \right]$$

Função objetivo:  $F(x) = \sum_{i=1}^n x_i^2$ ,  $\Omega^* = \{\mathbf{0}\}$ .

**Problema 24** DTLZ2-Res::

Problema multiobjetivo:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & (f_1(x), \dots, f_p(x)) \\ \text{s. a} \quad & \sum_{i=1}^n x_i \leq 1 \\ & x_i \in [0, 1] \quad \forall i = 1, \dots, n \end{aligned}$$

onde

$$\begin{aligned} n &\geq p, \\ f_1(x) &= (1 + g(x)) \prod_{i=1}^{p-1} \cos(x_i \pi / 2) \\ f_m(x) &= (1 + g(x)) \prod_{i=1}^{p-m} \cos(x_i \pi / 2) \sin(x_{p-m+1} \pi / 2), \quad m = 2, \dots, p-1 \\ f_p(x) &= (1 + g(x)) \sin(x_1 \pi / 2) \\ g(x) &= \sum_{i=p}^n (x_i - 0.5)^2 \end{aligned}$$

Função objetivo:  $F(x) = \sum_{i=1}^n x_i^2$ ,  $\Omega^* = \{\mathbf{0}\}$ .

**Problema 25** DTLZ3-Res:

Problema multiobjetivo:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & (f_1(x), \dots, f_p(x)) \\ \text{s. a} \quad & \sum_{i=1}^n x_i \leq 1 \\ & x_i \in [0, 1] \quad \forall i = 1, \dots, n \end{aligned}$$

onde

$$\begin{aligned} n &\geq p, \\ f_1(x) &= (1 + g(x)) \prod_{i=1}^{p-1} \cos(x_i \pi / 2) \\ f_m(x) &= (1 + g(x)) \prod_{i=1}^{p-m} \cos(x_i \pi / 2) \sin(x_{p-m+1} \pi / 2), \quad m = 2, \dots, p-1 \\ g(x) &= \sum_{i=p}^n (x_i - 0.5)^2 \end{aligned}$$

$$f_p(x) = (1 + g(x)) \operatorname{sen}(x_1\pi/2)$$
$$g(x) = 100 \left[ p + \sum_{i=p}^n (x_i - 0.5)^2 - \cos(20\pi(x_i - 0.5)) \right]$$

Função objetivo:  $F(x) = \sum_{i=1}^n x_i^2$ ,  $\Omega^* = \{\mathbf{0}\}$ .

---

---

## Capítulo 6

# Resultados dos experimentos

Neste capítulo apresentamos os resultados numéricos obtidos da implementação dos métodos propostos.

### 6.1 Introdução

Antes de mostrar as tabelas com os resultados, vamos discutir um problema que surgiu na implementação do problema 20. Também vamos expor a estratégia que usamos para contornar essa dificuldade o que nos leva a duas variantes na implementação dos métodos 1 e 2. Finalmente, com os resultados obtidos podemos concluir as vantagens de usar a nossa estratégia para problemas nos quais o nível inferior tem restrições de desigualdade.

### 6.2 RI sem perturbação ou RI com perturbação

No capítulo 1 analisamos o problema (1.3), para o qual achamos um ponto KKT na fase de otimização usando o método de RI, no qual é satisfeita a complementaridade estrita, mas um ponto ruim do ponto de vista da decisão original.

Esta discussão foi gerada porque surgiu uma dificuldade semelhante em um dos problemas implementados, o Problema 20, que vamos analisar a seguir. Reescrevendo-o temos:

**Problema 20:**

$$\begin{aligned} \min_{x \in \mathbb{R}^2} \quad & x_1^2 + x_2^2 \\ \text{s. a} \quad & x \in \mathcal{X}^* \end{aligned} \tag{6.1}$$

com  $\mathcal{X}^*$  o conjunto eficiente do problema multiobjetivo:

$$\begin{aligned}
 \min_{x \in \mathbb{R}^2} \quad & (f_1(x), \dots, f_p(x)) \\
 \text{s. a} \quad & x_2 + \frac{2}{5}x_1 \leq 2 \\
 & 5x_1 - x_2 \leq 10 \\
 & x_1, x_2 \geq 0
 \end{aligned} \tag{6.2}$$

onde

$$\begin{aligned}
 f_i(x_1, x_2) &= (x_1 - a_i)^2 + (x_2 - b_i)^2, \quad \forall i = 1, \dots, p \\
 a_i &= 2 \cos(\theta_i) + 1 \quad \forall i = 1, \dots, 2 \\
 b_i &= 2 \sin(\theta_i) + 1 \quad \forall i = 1, \dots, 2 \\
 \theta_i &= \frac{2\pi}{p}(i - 1)
 \end{aligned}$$

Vamos considerar o caso  $p = 3$ . Agregando variáveis de folga nas restrições de desigualdade e chamando-as de  $x_3$  e  $x_4$ , o problema multiobjetivo (6.2) fica:

$$\begin{aligned}
 \min_{x \in \mathbb{R}^4} \quad & (f_1(x), f_2(x), f_3(x)) \\
 \text{s. a} \quad & x_2 + \frac{2}{5}x_1 + x_3 = 2 \\
 & 5x_1 - x_2 + x_4 = 10 \\
 & x_1, x_2, x_3, x_4 \geq 0
 \end{aligned} \tag{6.3}$$

onde

$$\begin{aligned}
 f_i(x) &= (x_1 - a_i)^2 + (x_2 - b_i)^2, \quad \forall i = 1, \dots, 3 \\
 a_i &= 2 \cos(\theta_i) + 1 \quad \forall i = 1, \dots, 3 \\
 b_i &= 2 \sin(\theta_i) + 1 \quad \forall i = 1, \dots, 3 \\
 \theta_i &= \frac{2\pi}{3}(i - 1) \quad \forall i = 1, \dots, 3.
 \end{aligned}$$

O conjunto de pontos eficientes para este problema foi ilustrado na Figura 5.4.

Como foi desenvolvido no Capítulo 4, vamos transformar o problema (6.3) em um problema em dois níveis. Assim, dado  $\epsilon > 0$ , temos:

$$\begin{aligned}
& \min_{\alpha, x} && x_1^2 + x_2^2 \\
& s. a && \sum_{i=1}^3 \alpha_i = 1 \\
& && \alpha_i \geq \epsilon \quad \forall i = 1, \dots, 3 \\
& x = \arg \min_x && \alpha_1 f_1(x) + \alpha_2 f_2(x) + \alpha_3 f_3(x) \\
& s. a && x_2 + \frac{2}{5}x_1 + x_3 = 2 \\
& && 5x_1 - x_2 + x_4 = 10 \\
& && x_1, x_2, x_3, x_4 \geq 0
\end{aligned} \tag{6.4}$$

onde

$$\begin{aligned}
f_i(x) &= (x_1 - a_i)^2 + (x_2 - b_i)^2, \quad \forall i = 1, \dots, 3 \\
a_i &= 2 \cos(\theta_i) + 1 \quad \forall i = 1, \dots, 3 \\
b_i &= 2 \sin(\theta_i) + 1 \quad \forall i = 1, \dots, 3 \\
\theta_i &= \frac{2\pi}{3}(i - 1) \quad \forall i = 1, \dots, 3.
\end{aligned}$$

No método de Restauração Inexata para programação em dois níveis, a fase de otimalidade minimiza a função objetivo do primeiro nível sobre o espaço tangente às condições  $KKT(\alpha)$  do segundo nível no ponto obtido na fase de restauração (o qual chamamos de  $\pi(y^k)$  com  $y^k$  o ponto intermediário) e, uma região de confiança ao redor de  $y^k$ .

As condições  $KKT(\alpha)$  do segundo nível para o problema (6.4) estão dadas pelo seguinte sistema não linear:

$$\begin{aligned}
\sum_{i=1}^3 \alpha_i \nabla f_i(x) + \mu_1 \nabla h_1(x) + \mu_2 \nabla h_2(x) - \gamma &= 0 \\
h_1(x) &= 0 \\
h_2(x) &= 0 \\
x_1 \gamma_1 &= 0 \\
x_2 \gamma_2 &= 0 \\
x_3 \gamma_3 &= 0 \\
x_4 \gamma_4 &= 0 \\
x_i &\geq 0 \quad \forall i = 1, \dots, 4 \\
\gamma_i &\geq 0 \quad \forall i = 1, \dots, 4,
\end{aligned}$$

com  $h_1(x) = \frac{2}{5}x_1 + x_2 + x_3 - 2 = 0$ ,  $h_2(x) = 5x_1 - x_2 + x_4 - 10 = 0$ ,  $\mu = (\mu_1, \mu_2)$  o vetor dos multiplicadores associados às igualdades  $h_1(x) = 0$  e  $h_2(x) = 0$  respectivamente e  $\gamma = (\gamma_1, \gamma_2, \gamma_3, \gamma_4)$  o vetor dos multiplicadores das desigualdades  $x_i \geq 0$ ,  $i = 1, \dots, 4$ .

O espaço tangente é dado pelo seguinte sistema linear:  $C'(y^k)(v - y^k) = 0$  com  $v = (\alpha, x, \mu, \gamma)$  no conjunto viável, onde  $C'(y^k)$  representa a matriz jacobiana do sistema não linear que obtemos das condições  $KKT(\alpha)$  enunciadas acima, avaliadas no ponto intermediário.

A matriz jacobiana  $C'(\alpha, x, \mu, \gamma)$  para este problema no ponto intermediário  $y^k$  corresponde à seguinte matriz  $13 \times 10$ :

$$C'(y^k) = \begin{bmatrix} \nabla f_1(x^k) & \nabla f_2(x^k) & \nabla f_3(x^k) & \sum_{i=1}^3 \alpha_i^k \nabla^2 f_i(x^k) & \nabla h_1(x^k) & \nabla h_2(x^k) & I \\ 0 & \dots & 0 & \nabla h_1^T(x^k) & 0 & 0 & 0 \\ & & & \nabla h_2^T(x^k) & & & \\ \vdots & \ddots & \vdots & & \vdots & \vdots & \\ & & \text{diag}(\gamma_1^k, \dots, \gamma_4^k) & & & & \text{diag}(x_1^k, \dots, x_4^k) \\ 0 & \dots & 0 & & 0 & 0 & \end{bmatrix}$$

onde  $\text{diag}(\gamma_1^k, \dots, \gamma_4^k)$  e  $\text{diag}(x_1^k, \dots, x_4^k)$  correspondem a matrizes diagonais com  $\gamma_1^k, \dots, \gamma_4^k$  e  $x_1^k, \dots, x_4^k$  na diagonal respectivamente.

Neste problema, para resolver a fase de restauração começamos usando Restauração Exata, mas com certos pontos iniciais o ponto obtido na fase de restauração era o mesmo que o ponto obtido na fase de otimalidade, esse ponto continha:  $x_1^k = 0.6897$ ,  $x_2^k = 1.7241$

e  $x_3^k = 0$  com o multiplicador associado a este último valor  $\gamma_3^k > 0$ , como no problema (1.3). O que aconteceu foi o seguinte:

As equações da complementaridade estão dadas por:

$$x_1\gamma_1 = 0$$

$$x_2\gamma_2 = 0$$

$$x_3\gamma_3 = 0$$

$$x_4\gamma_4 = 0$$

A linearização destas equações, ou, as equações do espaço tangente no ponto  $y^k$  associadas à complementaridade são respectivamente:

$$\gamma_1^k x_1 + x_1^k \gamma_1 = 0$$

$$\gamma_2^k x_2 + x_2^k \gamma_2 = 0$$

$$\gamma_3^k x_3 + x_3^k \gamma_3 = 0$$

$$\gamma_4^k x_4 + x_4^k \gamma_4 = 0$$

substituindo o valor de  $x_3^k$  e  $\gamma_3^k$  obtido na fase de restauração, na terceira equação da linearização obtemos:

$$\gamma_3^k x_3 + 0 = 0,$$

portanto temos que  $x_3 = 0$ . Logo no conjunto viável para a otimalidade temos  $x_3 = 0$ , e a restrição linear

$$\frac{2}{5}x_1 + x_2 + x_3 = 2$$

fica

$$\frac{2}{5}x_1 + x_2 = 2.$$

Logo o ponto obtido na fase de otimalidade deve pertencer a essa reta e corresponde ao valor  $x_1 = 20/29 \approx 0.6897$ ,  $x_2 = 50/29 \approx 1.7241$ . De fato, isolando  $x_2$  da equação da reta temos  $x_2 = 2 - 2/5x_1$ , substituindo este valor na função objetivo e chamando de  $G(x_1)$  esta nova função temos  $G(x_1) = x_1^2 + (2 - 2/5x_1)^2$ . Logo  $G'(x_1) = 2x_1 - 4/5(2 - 2/5x_1)$ , da



condição de otimalidade de primeira ordem temos  $h'(x_1) = 0$ , isolando  $x_1$  desta equação temos  $x_1 = 20/29$  e portanto  $x_2 = 50/29$ . Da condição de otimalidade de segunda ordem temos que o ponto  $\hat{x} = (20/29, 50/29)$  corresponde a um mínimo local nessa reta.

Logo, o ponto obtido pelo algoritmo de RI na fase de otimalidade é um ponto KKT do subproblema dessa fase, mas não corresponde a um mínimo local do problema original, mais ainda, corresponde a um ponto ruim do ponto de vista da decisão original.

Assim, em um problema prático, bem geométrico e simples, surgiu a mesma dificuldade que no problema (1.3).

Qual foi a estratégia utilizada para contornar esta dificuldade? Usamos RI para resolver a fase de restauração, isto é, achamos o ponto intermediário inexatamente sem acrescentar folgas nas restrições de desigualdade, em seguida perturbamos as restrições de complementaridade como é feito no método de pontos interiores. O valor da perturbação é dado por:  $\zeta = x \cdot \gamma / n$ , onde  $x$  representa as variáveis primais e  $\gamma$  as variáveis duais do problema de segundo nível. Assim a complementaridade

$$x_i \gamma_i = 0 \quad \forall i = 1, \dots, 4$$

fica

$$x_i \gamma_i - \zeta = 0 \quad \forall i = 1, \dots, 4,$$

logo  $x_i^k \neq 0$  e  $\gamma_i^k \neq 0 \quad \forall i$ , e na linearização desta equação

$$\gamma_i^k x_i + x_i^k \gamma_i = 0 \quad \forall i,$$

os valores de  $x_i$  ou  $\gamma_i$  não ficam forçados a serem zero como aconteceu no problema antes mencionado.

Portanto, se na fase de restauração foram obtidos pontos estacionários do problema do primeiro nível, com a perturbação, a restrição de complementaridade não será satisfeita, e por isso haverá mais possibilidades desses pontos serem abandonados pelo algoritmo.

Pelo exposto acima, surgiram duas variantes para os Métodos 1 e 2:

**Variante 1:** Resolver a fase de restauração usando *Restauração Exata*, que corresponde

ao próprio método de *Restauração Inexata* e chamamos esta variante simplesmente de RI.

**Variante 2:** Resolver a fase de restauração usando *Restauração Inexata* e perturbar a restrição de complementaridade associada ao problema de nível inferior, chamamos esta variante de *Restauração Inexata Perturbada (RIP)*.

Estas duas variantes para os métodos 1 e 2 foram implementadas nas famílias de problemas com restrições de desigualdade além da caixa. Os resultados obtidos serão apresentados a seguir.

## 6.3 Resultados obtidos

Os testes computacionais foram realizados com o auxílio da linguagem **Matlab** R2012b (8.0.0.783) 64-bit (glnxa 64) e a rotina **fmincon.m** para resolver os subproblemas. Esta rotina é parte das ferramentas de otimização fornecidas pela linguagem, desenvolvida para resolver problemas de programação não linear. Dependendo da natureza do problema é possível escolher o algoritmo que melhor se adapte ao problema a resolver. Neste trabalho usamos as definições padrão estabelecidas pelo algoritmo mas, pela natureza dos problemas, o método escolhido pela própria **fmincon** foi *Programação Quadrática Sequencial* (SQP). A implementação do algoritmo SQP utilizado pela rotina **fmincon** baseia-se no Capítulo 18 de Nocedal e Wright [33]. Foi utilizada dupla precisão com os seguintes dados do computador: Intel(R), Core(TM) i3-3110M CPU @ 2.40 GHz  $\times$  4 com 3.8 GB de RAM .

A tolerância exigida para a medida da inviabilidade, assim como para a medida da otimalidade foi  $10^{-4}$ . Os parâmetros utilizados para o algoritmo foram:  $\delta_0 = 1$ ,  $r = 0.9$ ,  $\eta = 0.1$ ,  $\theta_{-1} = 0.5$ ,  $w_k = \frac{100}{(k+1)^{1.1}}$ ,  $\lambda_i^0 = 0 \ \forall \ i$  (multiplicador do Lagrangiano  $L(s, \lambda) = F(s) + C^T(s)\lambda$ ) e  $\zeta_k = x^k \cdot \gamma^k / n$ , com  $x^k, \gamma^k \in \Re^n$  as variáveis primais e duais do problema respectivamente na iteração  $k$ .

A seguir, são exibidos os resultados obtidos nestes testes. Eles são apresentados em diferentes tabelas cuja ordem de aparição corresponde à divisão dos grupos de problemas.

Para cada grupo, primeiro é apresentada a tabela correspondente ao Métodos 3 e em seguida as referentes aos Métodos 1 e 2. Foi escolhida esta ordem em virtude dos dados

apresentados em cada uma delas. Em todas as tabelas, na primeira coluna aparece o número do problema,  $\mathbf{p}$  representa a quantidade de coordenadas da função vetorial do problema multiobjetivo,  $\mathbf{n}$  a quantidade de variáveis do problema original. A notação usada dependendo da tabela é a seguinte:  $\alpha^0$  corresponde ao vetor inicial dos pesos para o problema multiobjetivo,  $\alpha^*$  representa o vetor solução dos pesos para o problema multiobjetivo,  $\mathbf{x}^0$  corresponde ao ponto inicial do problema original,  $\mathbf{x}^*$  é vetor solução do problema, *iter* quantidade de iterações que o algoritmo fez para convergir (Método 3), *it. ext*: iterações externas, corresponde às iterações feitas pelo algoritmo de *RI* externo, *it. int*: quantidade total de iterações internas realizadas pelo algoritmo (Método 1 e Método 2), aqui são somadas as iterações das fases de restauração, projeção ortogonal e otimalidade,  $\|d^{k,tan}\|$  norma infinito do gradiente projetado,  $\|C(s^*)\|$  norma infinito da restrição para o problema de primeiro nível. A notação  $(\cdot)^*$  significa que o ponto é um máximo do problema,  $(\cdot)^+$  significa que o ponto não representa um ponto propriamente eficiente. Nas tabelas referentes às famílias de problemas, denotaremos por *p.o.* o ponto ótimo do problema de primeiro nível, *p. KKT* os pontos *KKT* do primeiro nível que não representam o ótimo do problema do primeiro nível, e *p.efi* os pontos propriamente eficientes do segundo nível, mas que não correspondem a um ponto *KKT* do primeiro nível.

### 6.3.1 Testes preliminares

Os experimentos preliminares foram realizados com os problemas de 1 até 16 e de 20 até 22. Na seção de *Testes sistematizados*, apresentamos os resultado obtidos com as famílias de problemas 17, 18 e 19 (problemas multiobjetivo irrestritos) e as famílias de problemas 23, 24 e 25 (problemas multiobjetivo restritos), assim como os pontos iniciais escolhidos junto com as dimensões dos problemas rodados.

Prob.	$(p, n)$	$(\alpha^0, \mathbf{x}^0)$	$(\alpha^*, \mathbf{x}^*)$	$iter$	$F$	$\ C(s^*)\ $	Tempo
1	(2, 1)	(0.5, 0.5, -15)	(0.9999, 0.0001, 0.0001)	2	1.000E-04	2.9802E-08	0.0587
		(0.3, 0.7, 30)	(0.9999, 0.0001, 0.0001)	4	1.000E-04	2.9802E-08	0.0723
2	(3, 1)	(0.1, 0.3, 0.6, 40)	(0.9998, 0.0001, 0.0001, 0.0003)	4	3.000E-04	4.4499E-08	0.0886
		(0.05, 0.05, 0.9, -15)	(0.9998, 0.0001, 0.0001, 0.0003)	2	3.000E-04	2.9802E-08	0.0503
3	(2, 2)	(0.1, 0.9, -1, 75)	(0.9998, 0.0001, 0.0001, 0.0001)	3	2.000E-04	0.2980E-07	0.0675
		(0.5, 0.5, 0.1, 0.5)	(0.9999, 0.0001, 0.0001, 0.0001)	2	2.000E-04	0.2980E-07	0.0413
		(0.2, 0.8, -10, 5)	(0.9999, 0.0001, 0.0001, 0.0001)	2	2.000E-04	0.2980E-07	0.0442
4	(4, 2)	(0.6, 0.6, 0.6, 0.6, 10, 5)	(0.0001, 0.0001, 0.0001, 0.9997, 1, -1.9992)	5	-0.9992	0.5296 E-07	0.1105
		(0.8, 0.8, 0.1, 0.1, -4, -10)	(0.0001, 0.0001, 0.0001, 0.9997, 1, -1.9992)	3	-0.9992	0.9658 E-07	0.0773
		(0.1, 0.9, 0.9, 0.1, 2, 10)	(0.0001, 0.0001, 0.0001, 0.9997, 1, -1.9992)	5	-0.9992	0.6025 E-07	0.0964
5	(2, 2)	(0.1, 0.9, -5, 9)	(0.9999, 0.0001, 0.0001, -0.0001)	2	-1.7764E-15	0.2980 E-07	0.0449
		(0.6, 0.1, 10, 10)	(0.8888, 0.1112, 0.1112, -0.1112)	3	-8.3320E-08	0.1838 E-06	0.0477
		(0.01, 0.03, -5, 9)	(0.9999, 0.0001, 0.0001, -0.0001)	5	4.8943E-12	0.2980 E-07	0.0944
		(0.9, 0.3, 13, 9)	(0.0001, 0.9999, 0.9999, -0.9999)	3	2.5472E-08	0.2980 E-08	0.0563
6	(4, 2)	(0.1, 0.9, 0.1, 0.1, 10, 10)	(0.5197, 0.0001, 0.0001, 0.4801, 0.4804, -0.4799)	5	5.000E-04	0.32799 E-07	0.1403
		(0.1, 0.9, 0.9, 0.1, 10, 10)	(0.4724, 0.0001, 0.0001, 0.5274, 0.5277, -0.5272)	5	5.0000E-04	0.3015 E-07	0.1050
		(1, 0, 0, 0, 10, 10)	(0.0243, 0.0001, 0.0001, 0.9755, 0.9858, -0.9753)	3	5.0000E-04	0.4601 E-07	0.0794
		(0.6, 0, 0, 0.4, 10, 10)	(0.0162, 0.0001, 0.0001, 0.9836, 0.9839, -0.9834)	4	5.0000E-04	0.1399 E-07	0.0849
		(0.6, 0, 0, 0.6, 10, 10)	(0.9997, 0.0001, 0.0001, 0.0001, 0.0004, 0.0001)	6	5.0000E-04	0.3562 E-07	0.1311
7	(5, 3)	(0.1, 0.1, 0.1, 0.1, 0.1, 10, 10, 10)	(0.0001, 0.0001, 0.0001, 0.0001, 0.9996, -0.0001, -0.0001, 0.9996)	6	1.6E-07	0.2675 E-07	0.1637
		(0.9, 0.9, 0.9, 0.1, 0.1, -15, 0, 10)	(0.0001, 0.0001, 0.0001, 0.0001, 0.9996, -0.0001, -0.0001, 0.9996)	5	1.6E-07	0.2980 E-07	0.1389
8	(5, 3)	(0.1, 0.1, 0.1, 0.1, 0.9, 10, 10, 10)	(0.4885, 0.4885, 0.0115, 0.0115, 0.0001, 0.4770, 0.4770, 0.0001)	3	9.9965E-05	0.5969 E-07	0.0907
		(0.9, 0.9, 0.9, 0.1, 0.1, -15, 10, 10)	(0.0001, 0.9996, 0.0001, 0.0001, 0.0001, 0, 0.9995, 0.0001)	3	1.0007E-04	0.1736E-06	0.1121
		(0, 0, 0.1, 0.1, 0, -15, 10, 10)	(0.0001, 0.0001, 0.9996, 0.0001, 0.0001, -0.9995, 0, 0.0001)	4	1.0003E-04	0.9318E-07	0.1173
		(0, 0.9, 0.9, 0, 0.9, -15, -10, 10)	(0.0001, 0.0001, 0.9996, 0.0001, 0.0001, -0.9995, 0, 0.0001)	3	9.9947E-05	0.1295E-06	0.1045

Tabela 6.1: Problemas sem restrições - Método 3

Prob.	$(p, n)$	$(\alpha^*, \mathbf{x}^*)$	$it. \ ext.$	$it. \ int.$	$F$	$\ d^{k, tan}\ $	$\ C(s^*)\ $	Tempo
1	(2, 1)	(0.9999, 0.0001, 0.0001)	2	7	1.0013E-04	0	2.5431E-07	0.2916
		(0.9999, 0.0001, 0.0001)	2	8	9.9619E-05	0	7.6290E-07	0.2895
2	(3, 1)	(0.9998, 0.0001, 0.0001, 0.0003)	3	9	2.9962E-04	1.1102E-16	5.9921E-04	0.6481
		(0.9998, 0.0001, 0.0001, 0.0003)	3	9	3.0013E-04	1.55045E-17	6.0022E-04	0.6787
3	(2, 2)	(0.9999, 0, 0001, 0.0001, 0.0001)	2	9	1.9995E-04	0	5.9605E-08	0.3887
		(0.9999, 0, 0001, 0.0001, 0.0001)	2	8	1.9999E-04	0	1.4901E-08	0.3414
		(0.9999, 0, 0001, 0.0001, 0.0001)	2	8	2.0010E-04	0	1.9073E-07	0.3604
4	(4, 2)	(0.0001, 0.0001, 0.0001, 0.9997, 1, -1.9992)	2	12	-0.9992	8.0353E-17	3.1423E-08	0.5341
		(0.0001, 0.0001, 0.0001, 0.9997, 1, -1.9992)	3	18	-0.9992	2.2204E-16	2.6472E-08	0.8584
		(0.0001, 0.0001, 0.0001, 0.9997, 1, -1.9992)	3	21	-0.9992	1.1018E-17	3.9228E-08	0.8861
5	(2, 2)	(0.1, 0.9, 0.9, -0.9)	1	2	-4.2386E-08	0	8.4771E-08	0.1085
		(0.8219, 0.1781, 0.1781, -0.1781)	2	8	-8.5714.0055E-10	0	1.4880E-08	0.3296
		(0.2504, 0.7496, 0.7496, -0.7496)	2	9	3.2359E-09	0	1.6388E-08	0.3329
		(0.7795, 0.2205, 0.2205, -0.2205)	2	8	-2.4883E-09	0	1.9296E-08	0.3967
6	(4, 2)	(0.6536, 0.0001, 0.0001, 0.3462, 0.3465, -0.3460)	2	11	5.0000E-04	1.1102E-16	1.2523E-08	0.4689
		(0.5449, 0.0001, 0.0001, 0.4549, 0.4552, -0.4547)	2	10	4.9999E-04	6.6529E-17	1.7657E-08	0.4641
		(0.9997, 0.0001, 0.0001, 0.0001, 0.0004, 0.0001)	2	7	4.9981E-04	0	1.9073E-07	0.3876
		(0.5999, 0.0001, 0.0001, 0.3999, 0.4002, -0.3997)	2	9	5.0000E-04	0	1.2263E-07	0.4385
		(0.4999, 0.0001, 0.0001, 0.4999, 0.5002, -0.4998)	2	10	4.1672E-04	0	9.3177E-05	0.4386
7	(5, 3)	(0.0001, 0.0001, 0.0001, 0.0001, 0.9996, 0.0001, 0.0001, 0.9996)	3	14	1.6000E-07	6.0986E-19	2.2352E-08	0.9447
		(0.0001, 0.0001, 0.0001, 0.0001, 0.9996, 0.0000, 0.0000, 0.9996)	2	12	1.6001E-07	1.1018E-17	1.6570E-08	0.5791
8	(5, 3)	(0.25, 0.25, 0.25, 0.25, 0.0001, 0.0000, 0.0000, 0.0001)	2	10	9.9991E-05	5.5511E-17	1.92992E-08	0.1406
		(0.3354, 0.3291, 0.3354, 0.0001, 0.0001, 0.9995, 0.3290, 0.0001)	2	13	1.0001E-04	2.5697E-16	1.6270E-08	0.1438
		(0.0039, 0.0039, 0.4961, 0.4961, 0.0001, -0.4922, -0.4922, 0.0001)	2	10	1.0001E-04	5.5511E-17	1.7876E-08	0.1531
		(0.0416, 0.4583, 0.4583, 0.0416, 0.0001, -0.4167, 0.4167, 0.0001)	2	10	9.9994E-05	5.5511E-17	1.7190E-08	0.1312

Tabela 6.2: Problemas sem restrições - Método 1

Prob.	$(p, n)$	$(\alpha^*, \mathbf{x}^*)$	<i>it. ext.</i>	<i>it. int.</i>	$F$	$\ d^{k, \tan}\ $	$\ C(s^*)\ $	Tempo
1	(2, 1)	(0.9999, 0.0001, 0.0001)	2	7	1.0013E-04	0	2.5431E-07	0.2427
		(0.9999, 0.0001, 0.0001)	2	8	9.9619E-05	0	7.6290E-07	0.2398
2	(3, 1)	(0.9998, 0.0001, 0.0001, 0.0003)	3	15	2.9962E-04	2.9956E-05	7.6294E-07	0.4760
		(0.9998, 0.0001, 0.0001, 0.0003)	3	14	3.0013E-04	3.0005E-05	2.5431E-07	0.4415
3	(2, 2)	(0.9999, 0, 0001, 0.0001, 0.0001)	2	9	1.9995E-04	0	5.9605E-08	0.2869
		(0.9999, 0.0001, 0.0001, 0.0001)	2	8	1.9999E-04	0	1.4901E-08	0.2833
		(0.9999, 0.0001, 0.0001, 0.0001)	2	8	2.0010E-04	0	1.9073E-07	0.2723
4	(4, 2)	(0.0001, 0.0001, 0.0001, 0.9997, 1, -1.9992)	2	12	-0.9992	8.0353E-17	3.1423E-08	0.3516
		(0.0001, 0.0001, 0.0001, 0.9997, 1, -1.9992)	3	18	-0.9992	2.2204E-16	2.6472E-08	0.6094
		(0.0001, 0.0001, 0.0001, 0.9997, 1, -1.9992)	3	21	-0.9992	1.1018E-17	3.9228E-08	0.6680
5	(2, 2)	(0.1, 0.9, 0.9, -0.9)	1	2	-4.2386E-08	0	8.4771E-08	0.1603
		(0.8219, 0.1781, 0.1781, -0.1781)	2	8	-8.5714E-10	0	1.4880E-08	0.2205
		(0.2504, 0.7496, 0.7496, -0.7496)	2	9	3.2359E-09	0	1.6388E-08	0.3028
		(0.7795, 0.2205, 0.2205, -0.2205)	2	8	-2.4883e-09	0	1.9296E-08	0.2828
6	(4, 2)	(0.6536, 0.0001, 0.0001, 0.3462, 0.3465, -0.3460)	2	11	5.0000E-04	1.1235E-17	1.2523E-08	0.3607
		(0.5449, 0.0001, 0.0001, 0.4549, 0.4552, -0.4547)	2	10	4.9999E-04	6.6529E-17	1.7657E-08	0.3699
		(0.9997, 0.0001, 0.0001, 0.0001, 0.0004, 0.0001)	2	7	4.9981E-04	0	1.9073E-07	0.2617
		(0.5999, 0.0001, 0.0001, 0.3999, 0.4002, -0.3997)	2	9	5.0002E-04	0	1.2263E-07	0.3093
		(0.4999, 0.0001, 0.0001, 0.4999, 0.5002, -0.4998)	2	10	4.1672E-04	0	9.3177E-05	0.3638
7	(5, 3)	(0.0001, 0.0001, 0.0001, 0.0001, 0.9996, 0.0000, 0.0000, 0.9996)	3	14	1.6000E-07	6.0986E-19	2.2352E-08	0.6014
		(0.0001, 0.0001, 0.0001, 0.0001, 0.9996, 0.0000, 0.0000, 0.9996)	2	12	1.6001E-07	1.1018E-17	1.6570E-08	0.3642
8	(5, 3)	(0.25, 0.25, 0.25, 0.25, 0.0001, 0, 0, 0.0001)	2	10	9.9991E-05	5.5511E-17	1.9292E-08	0.3982
		(0.3354, 0.3291, 0.3354, 0.0001, 0.0001, -0.0000, 0.3290, 0.0001)	2	10	1.0004E-04	0	1.4875E-08	0.3674
		(0.0039, 0.0039, 0.4961, 0.4961, 0.0001, -0.4922, -0.4922, 0.0001)	2	10	1.0001E-04	5.5511E-17	1.7876E-08	0.3490
		(0.0416, 0.4583, 0.4583, 0.0416, 0.0001, -0.4168, 0.4167, 0.0001)	2	10	9.9992E-05	5.5511E-17	1.71902E-08	0.3657

Tabela 6.3: Problemas sem restrições - Método 2

Prob.	$(p, n)$	$(\alpha^0, \mathbf{x}^0)$	$(\alpha^*, \mathbf{x}^*)$	$iter$	$F$	$\ C(s^*)\ $	Tempo
9-ZDT1 $n1$	$(2, 1)$	$(0.72, 0.28, 10)$	$(0.3333, 0.6667, 0.9999)$	3	9.2908E-09	0.3740E-04	0.0658
		$(0.55, 0.45, 0.0001)$	$(0.9665, 0.0335, 0.0003)^*$	4	0.9994	0.2440E-04	0.0984
		$(0.25, 0.75, 40)$	$(0.3077, 0.6923, 1.0000)$	2	0	0.1490E-07	0.0534
9-ZDT1 $n2$	$(2, 2)$	$(0.9, 0.1, 20, 10)$	Não converge	—	—	—	—
		$(0.45, 0.55, 10, 40)$	Não converge	—	—	—	—
		$(0.82, 0.18, 0.0001, 10)$	Não converge	—	—	—	—
10	$(2, 2)$	$(0.8, 0.2, -2, 20)$	$(0.7285, 0.2715, 1.0000, 1.0000)^*$	7	2	0.2968 E-07	0.1853
		$(0.4, 0.6, -10, 10)$	$(0.3790, 0.6210, 1.0000, 1.0000)^*$	7	2	0.2978 E-04	0.1461
11	$(2, 2)$	$(0.5, 0.5, -10, 10)$	$(0.5114, 0.4886, 1.0000, 1.0000)$	7	2.2278E-22	0.2978 E-07	0.1535
		$(0.4, 0.6, -10, 10)$	$(0.4405, 0.5595, 1.0000, 1.0000)$	7	2.2164E-22	0.2978E-07	0.1376
12	$(2, 2)$	$(0.5, 0.5, -10, 10)$	$(0.5, 0.5, 0, 0)^*$	2	2	0.4470 E-07	0.0527
		$(0.1, 0.9, 40, 100)$	$(0.5, 0.5, 0.0005, 0)^*$	17	1.9991	0.1490 E-07	0.3258
13	$(2, 2)$	$(0.5, 0.5, -10, 10)$	$(0.16667, 0.8333, 2, 0)^+$	3	6.2500	0.1759 E-07	0.0741
		$(0.73, 0.27, 10, 20)$	$(0.0001, 0.9999, 4.4982, 0.0001)$	6	3.2509E-06	0.9959 E-07	0.1054
		$(0.23, 0.77, 10, 20)$	$(0.0001, 0.9999, 4.4982, 0.0001)$	7	3.2637E-06	0.5531 E-05	0.1237

Tabela 6.4: Problemas com restrições - Método 3

Prob.	$(p, n)$	$(\alpha^*, \mathbf{x}^*)$	$it. \ ext.$	$it. \ int.$	$F$	$\ d^{k,tan}\ $	$\ C(s^*)\ $	Tempo
9-ZDT1 $n1$	(2, 1)	(0.0001, 0.9999, 1)	2	15	0	0	1.5138E-13	0.6320
		(0.0001, 0.9999, 1)	3	20	0	1.1102E-16	1.5132E-13	0.8117
		(0.2500, 0.7500, 1)	1	1	0	0	0	0.0669
9-ZDT1 $n2$	(2, 2)	(0.1771, 0.8229, 1, 0)	5	46	0	1.9429E-16	1.2278E-07	2.2911
		(0.24230, 0.7577, 1, 0)	2	13	5.4530E-17	0	1.1368E-07	0.5928
		(0.2607, 0.7393, 1, 0)	4	40	0	0	1.1125E-07	1.7553
10	(2, 2)	0.4792, 0.5208, 0, 0)	4	43	1.8129E-36	0	7.7612E-09	2.2749
		(0.4, 0.6, 0, 0)	1	10	9.4885E-36	0	8.9411E-09	0.2895
11	(2, 2)	(0.8331, 0.1669, 0.9994, 0.9988)	9	117	1.7460E-07	9.7750E-05	2.0707E-07	3.5833
		(0.4, 0.6, 0, 0)*	1	11	2	0	8.9411E-09	0.2349
12	(2, 2)	(0.5, 0.5, 0, 0)*	1	4	2	3.9611E-16	5.7233E-08	0.2153
		(0.0001, 0.9999, 0.6299, 0.3968)	2	21	0.5008	0	4.0444E-05	0.8395
13	(2, 2)	(0.0001, 0.9999, 4.4982, 0.0001)	6	51	3.2565E-19	5.5965E-19	2.9079E-04	2.5300
		(0.0001, 0.9999, 4.4982, 0.0001)	8	70	3.2749E-06	2.7397E-16	2.9147E-05	3.3486
		(0.0001, 0.9999, 4.4982, 0.0001)	5	48	3.2398E-06	6.7755E-06	2.9024E-04	2.0261

Tabela 6.5: Problemas com restrições - Método 1

Prob.	$(p, n)$	$(\alpha^*, \mathbf{x}^*)$	$iter.$	$it. \ int.$	$F$	$\ d^{k,tan}\ $	$\ C(s^*)\ $	Tempo
9-ZDT1 $n1$	(2, 1)	(0.0001, 0.9999, 1)	2	28	0	0	1.5138E-13	0.3734
		(0.0001, 0.9999, 1)	2	22	0	0	1.5132E-13	0.2956
		(0.25, 0.75, 1)	1	2	0	0	0	0.0608
9-ZDT1 $n2$	(2, 2)	(0.1771, 0.8229, 1, 0)	5	46	0	0	1.2263E-07	1.0258
		(0.2423, 0.7577, 1, 0)	2	13	5.4530E-17	0	1.1368E-07	0.3418
		(0.2607, 0.7393, 1, 0)	1	38	1.2326E-32	0	1.1212E-07	0.8268
10	(2, 2)	(0.4792, 0.5208, 0, 0)	4	40	1.4130E-35	0	7.8127E-09	1.3191
		(0.4, 0.6, 0, 0)	1	10	7.1909E-36	0	8.9385E-09	0.2276
11	(2, 2)	(0.8376, 0.1624, 1, 1)	2	15	1.2252E-22	0	4.8744E-07	0.4981
		(0.4, 0.6, 0, 0)*	1	11	2	0	8.9422E-09	0.2946
12	(2, 2)	(0.5, 0.5, 0.000, 0.0000)*	1	4	2	3.961E-16	5.7233E-08	0.1766
		(0.0001, 0.9999, 0.6299, 0.3968)	2	19	0.5008	0	9.6126E-06	0.5710
13	(2, 2)	(0.0001, 0.9999, 4.4982, 0.0001)	6	51	3.2519E-06	5.5976E-19	2.9079E-04	1.7293
		(0.0001, 0.9999, 4.4982, 0.0001)	7	63	3.2842E-06	1.1102E-16	2.9171E-04	2.0884
		(0.0001, 0.9999, 4.4982, 0.0001)	5	48	3.2398E-06	6.7575E-19	2.9024E-04	1.3835

Tabela 6.6: Problemas com restrições - Método 2

Das Tabelas 6.1 a 6.3, podemos dizer que os três métodos chegam à solução do problema, mas o método que precisa menor tempo até convergir é o Método 3. Nas seguintes três Tabelas 6.4 a 6.6 já notamos diferença nos resultados obtidos para os diferentes métodos. Para o problema que não é convexo nem diferenciável, no caso bidimensional, a reformulação KKT não consegue obter ponto viável, já os Métodos 1 e 2 chegam ao ótimo



do problema.

Prob.	$(p, n)$	$(\alpha^0, \mathbf{x}^0)$	$\mathbf{x}^*$	$iter$	$F$	$\ C(s^*)\ $	Tempo
14	(2, 2)	(0.75, 0.25, 10, 10)	<i>p.o.</i>	3	0	0.1490E-07	0.6634
		(0.35, 0.65, -10, -10)	<i>p.o.</i>	3	0	0	0.3295
		(0.5, 0.5, 0.5, 0.5)	<i>p. KKT</i>	11	0.1604	0.5328E-10	1.0344
	(6, 6)	$\alpha_i = 1/6, x_i = 0.5$	<i>p. KKT</i>	37	0.1126	0.1079E-08	19.7792
	(40, 40)	$\alpha_i = 1/40, x_i = 0.5$	<i>p.o.</i>	3	0	0	831.6970
15	(2, 2)	(0.5, 0.5, 0.7, 0.7)	<i>p.o.</i>	16	2.6286E-08	0.8131E-19	1.4781
		(0.75, 0.25, 10, 10)	<i>p. KKT</i>	3	1	0.1111E-03	0.3512
		(0.25, 0.75, 10, 10)	<i>p. KKT</i>	3	1	0.1111E-03	0.5423
	(6, 6)	$\alpha_i = 1/6, x_i = 0.7$	não <i>p.efi.</i>	22	5.9000E-04	0.4295E-08	11.0648
	(50, 50)	$\alpha_i = 1/50, x_i = 0.5$	<i>p.o.</i>	9	0.0016	0.1E-07	6.8198E+03
16	(2, 2)	(0.6, 0.4, 0.3, 10)	<i>p.o.</i>	7	4.0537E-09	0.2711E-19	0.8260
		(0.5, 0.5, -10, 10)	<i>p.o.</i>	4	0	0.4718E-15	0.4439
	(6, 6)	$\alpha_i = 1/6, x_i = 0.5$	<i>p.o.</i>	3	0	0.1388E-16	1.4922
	(12, 12)	$\alpha_i = 1/12, x_{i=1,\dots,6} = 10, x_{i=7,\dots,12} = 0$	<i>p.o.</i>	6	4.8923E-07	0	9.0605
		$\alpha_{i=1,\dots,11} = 0.7/11, \alpha_{12} = 0.3, x_i = 0.3$	<i>p.o.</i>	4	6.4044E-08	0	5.4252
		$\alpha_1 = 0.3, x_1 = 0.3, \alpha_{i=2,\dots,12} = 0.7/11$	<i>p.o.</i>	70	6.2593E-07	0	98.5023
	(50, 50)	$\alpha_i = 1/50, x_i = 0.8$	<i>p.o.</i>	95	3.4895E-07	0	3.6388E+03

**Tabela 6.7:** Família de problemas sem restrições - Método 3

Prob.	$(p, n)$	$(\alpha^*, \mathbf{x}^*)$	$it. \ ext.$	$it. \ int.$	$F$	$\ d^{k,tan}\ $	$\ C(s^*)\ $	Tempo
14	(2, 2)	$p. \ KKT$	1	4	1	0	5.5355E-10	0.2336
		$p.o.$	1	3	0	0	0	0.0950
		$p. \ KKT$	1	5	1	0	9.5651E-09	0.2561
	(6, 6)	$p. \ KKT$	1	25	5	0	1.9868E-08	2.8004
	(40, 40)	$p. \ KKT$	2	34	5	6.9389E-32	4.0976E-08	1.0914E+03
15	(2, 2)	$p.o.$	2	13	2.6283E-08	0	9.2674E-09	1.1102
		$p.o.$	2	13	2.6283E-08	0	1.6185E-08	1.1847
		$p.o.$	2	15	2.6284E-08	0	8.7620E-09	1.2131
	(6, 6)	$p. \ KKT$	3	39	1	3.8165E-05	3.0659E-08	7.2705
	(50, 50)	$p. \ KKT$	3	47	2	4.0078E-05	1.9092E-07	3.7518E+03
16	(2, 2)	$p.o.$	1	4	0	0	1.8477E-08	0.2153
		$p.o.$	1	3	0	0	7.9466E-08	0.1657
	(6, 6)	$p.o.$	2	10	0	0	1.3412E-08	1.7581
	(12, 12)	$p.o.$	1	19	0	0	1.9536E-08	4.7234
		$p. \ KKT$	2	12	1	2.7756E-17	5.1811E-08	5.3611
		$p.o.$	2	12	0	0	1.4106E-08	5.4934
	(50, 50)	$p.o.$	2	18	0	0	2.8688E-08	89.3994

Tabela 6.8: Família de problemas sem restrições - Método 1

Prob.	$(p, n)$	$\mathbf{x}^*$	$it. \ ext.$	$it. \ int.$	$F$	$\ d^{k,tan}\ $	$\ C(s^*)\ $	Tempo
14	(2, 2)	$p. \ KKT$	1	4	1	0	5.5355E-10	0.2011
		$p.o.$	1	1	0	0	0	0.0874
		$p. \ KKT$	1	5	1	0	9.5651E-09	0.2041
	(6, 6)	$p. \ KKT$	1	25	5	0	1.9868E-08	2.8470
	(40, 40)	$p. \ KKT$	2	34	5	6.9389E-18	4.0976E-08	241.4191
15	(2, 2)	$p.o.$	2	13	2.6283E-08	0	9.2674E-09	0.7347
		$p.o.$	2	13	2.6286E-08	0	1.6185E-08	0.6744
		$p.o.$	2	15	2.6284E-08	2.6662E-12	8.7620E-09	0.7323
	(6, 6)	$p. \ KKT$	3	41	1	3.8165E-05	3.0659E-08	6.5207
	(50, 50)	$p. \ KKT$	3	49	2	4.0058E-05	2.9675E-07	856.3584
16	(2, 2)	$p.o.$	1	4	0	0	1.8477E-08	0.1412
		$p.o.$	1	3	0	0	7.9466E-08	0.1164
	(6, 6)	$p.o.$	2	10	0	0	1.3412E-08	1.2563
	(12, 12)	$p.o.$	2	20	0	0	1.9536E-08	5.2835
		$p. \ KKT$	2	12	1	0	5.1811E-08	4.4488
		$p.o.$	2	12	0	0	1.4106E-08	3.6125
	(50, 50)	$p.o.$	2	18	0	0	1.3787E-08	68.5324

Tabela 6.9: Família de problemas sem restrições - Método 2

Prob.	$(p, n)$	$(\alpha^0, \mathbf{x}^0)$	$\mathbf{x}^*$	$iter$	$F$	$\ C(s^*)\ $	Tempo
20	(3, 2)	(0.2, 0.3, 0.5, 10, 10)	<i>p.e fi.</i>	5	1.6065	0.0403E-10	1.6065
	(3, 2)	(0.2497, 0.6272, 0.1231, 0.6897, 1.7241)	<i>p.o.</i>	10	9.0000e-08	0.2007 E-09	1.9863
	(3, 2)	(0.3, 0.6, 0.1, 0.6897, 1.7241)	<i>p.o.</i>	12	9.0000e-08	0.1011 E-09	1.4823
	(6, 2)	$\alpha_i = 1/6, x_i = -10$	<i>p.e fi.</i>	10	7.1106E-06	1.000E-11	3.0891
	(100, 2)	$\alpha_i = 1/100, x_i = 0.9$	<i>p.e fi.</i>	3	3.6377	0.2129E-12	73.5020
	(300, 2)	$\alpha_i = 1/300, x_1 = 10, x_2 = 1$	<i>p.e fi.</i>	3	2.4323	0.6343E-12	636.4001
21	(6, 6)	$\alpha_i = 1/6, x_i = 0.3$	não factível	1	0.54	0.7288	2.0106
	(10, 10)	$\alpha_i = 1/10, x_{i=1,\dots,3} = 0.4, x_{i=4,\dots,10} = 0$	não <i>KKT</i>	76	0.2524	0.5168E-06	135.2789
		$\alpha_i = 1/10, x_{i=1,\dots,5} = 0.4, x_{i=6,\dots,10} = 0$	não <i>KKT</i>	38	0.5394	0.2326E-06	77.7624
	(20, 20)	$\alpha_i = 1/20, x_{i=1,\dots,10} = 0.4, x_{i=11,\dots,20} = 0$	não <i>KKT</i>	11	0.1503	0.1848E-04	273.7873
22	(2, 2)	(0.35, 0.65, -0.3, -0.3))	não factível	38	0.5521	0.3930E-04	10.3718
		(0.85, 0.15, 0, 0))	<i>p.KKT</i>	19	0.5	0.3331 E-15	2.3496
	(3, 3)	$\alpha_i = 1/3, x_{i=1,\dots,3} = 0.5$	<i>p.KKT</i>	15	0.5202	0.8838E-10	2.4982
	(6, 6)	$\alpha_i = 1/6, x_{i=1,\dots,6} = 0.3$	não factível	1	0.5400	0.7288	1.8058
	(20, 20)	$\alpha_i = 1/20, x_{i=1,\dots,20} = 0.5$	não <i>p.e fi.</i>	88	0.0577	0.3280 E-07	1.8003E+03

Tabela 6.10: Família de problemas com restrições - Método 3

PROBLEMAS USANDO RI-LAGRANGIANO								
Prob.	$(p, n)$	$(\alpha^*, \mathbf{x}^*)$	$it. \ ext.$	$it. \ int.$	$F$	$\ d^{k,tan}\ $	$\ C(s^*)\ $	Tempo
20	(3, 2)	$p.o.$	4	69	9.0162E-08	2.5197E-16	5.3996E-07	19.4553
		$p.o.$	4	71	9.0229E-08	2.0771E-16	7.6134E-07	18.6243
		$p.KKT$	2	437	3.4483	1.4107E-07	2.5927E-07	107.6615
	(6, 2)	$p.o.$	5	64	6.2739E-18	0	8.5716E-07	44.4248
	(100, 2)	$p.o.$	5	160	2.0981E-13	1.1606E-08	8.2717E-07	772.0355
	(300, 2)	$p.o.$	5	243	8.4798E-34	0	6.6029E-07	6.8420E+03
21	(6, 6)	$p. \ max.$	4	216	1	0	1.8875E-08	114.1845
	(10, 10)	$p. \ max.$	2	55	1	0	7.0243E-08	52.8018
		$p.o.$	2	12	1.6319E-31	0	1.7881E-08	10.9889
	(20, 20)	$p.o.$	2	13	0	0	2.0862E-08	117.7985
22	(2, 2)	$p.o.$	1	6	1	0	1.5636E-08	0.2931
	(2, 2)	$p.o.$	12	293	1	0	4.0275E-09	29.9104
		$p.KKT$	11	350	1	0	1.9209E-09	35.9664
	(6, 6)	$p. \ KKT$	16	359	0.25	0	1.6516E-09	87.0860
	(20, 20)	$p. \ KKT.$	5	105	0.0556	8.4143E-05	4.1327E-09	311.2713
PROBLEMAS PERTURBADOS USANDO RIP-LAGRANGIANO								
Prob.	$(p, n)$	$(\alpha^*, \mathbf{x}^*)$	$it. \ ext.$	$it. \ int.$	$F$	$\ d^{k,tan}\ $	$\ C(s^*)\ $	Tempo
20	(3, 2)	$p.o.$	2	13	9.0003E-08	3.1101E-16	8.9382E-09	1.7118
		$p.o.$	4	30	9.0007E-08	1.1102E-16	1.2645E-07	4.5602
		$p.o.$	4	32	8.9998E-08	1.1102E-16	2.7977E-08	4.4218
	(6, 2)	$p.o.$	2	15	1.2338E-15	0	2.2328E-07	2.3561
	(100, 2)	$p.o.$	2	9	0	0	7.7715E-08	21.7779
	(300, 2)	$p.o.$	2	10	1.3034E-14	9.1683E-17	1.9040E-07	89.0647
21	(6, 6)	$p. \ o.$	1	3	0	0	9.9341E-09	0.7401
	(10, 10)	$p. \ max.$	2	7	1	1.3878E-17	5.1011E-09	1.8351
		$p.o.$	1	3	0	0	1.7881E-08	1.1206
	(20, 20)	$p.o.$	1	6	1.0000E-07	2.0007E-05	1.1016E-08	6.1333
22	(2, 2)	$p.o.$	1	8	1	0	1.5636E-08	0.6293
	(2, 2)	$p.o.$	3	86	1	0	6.881E-09	7.5433
		$p. \ KKT$	1	9	2	0	4.9671E-09	0.6742
	(6, 6)	$p. \ KKT$	1	28	5	0	1.9868E-08	3.0980
	(20, 20)	$p. \ KKT.$	2	38	5	1.3715E-16	4.5657E-08	69.7157

Tabela 6.11: Família de problemas com restrições - Método 1

PROBLEMAS USANDO RI-FUNÇÃO $F(x)$								
Prob.	$(p, n)$	$(\alpha^*, \mathbf{x}^*)$	$it. \ ext.$	$it. \ int.$	$F$	$\ d^{k,tan}\ $	$\ C(s^*)\ $	Tempo
20	(3, 2)	$p.o.$	4	77	1.2013E-07	8.0282E-06	6.5191E-09	18.2797
		$p.KKT$	2	42	3.4483	4.9152E-06	3.7422E-07	14.0259
		$p.KKT$	2	39	3.4483	0	2.4596E-07	14.7668
	(6, 2)	$p.o.$	2	63	1.2532E-15	0	8.5716E-07	27.9219
	(100, 2)	$p.o.$	5	136	3.2279E-07	5.1423E-05	9.6590E-08	449.2247
	(300, 2)	$p.o.$	5	289	2.8402E-13	1.4416E-08	6.6029E-07	3.1338E+03
21	(6, 6)	$p. \ max.$	4	154	1	0	7.6102E-08	10.9401
	(10, 10)	$p. \ max.$	2	58	1	0	7.0243E-08	7.7445
		$p.o.$	2	12	1.6319E-31	0	1.7881E-08	2.6001
	(20, 20)	$p.o.$	2	16	5.3226E-14	1.3878E-17	3.3907E-08	11.5957
22	(2, 2)	$p.o.$	1	6	1	0	1.5636E-08	0.3126
	(2, 2)	$p.o.$	12	293	1	0	4.0275E-09	16.0528
	(3, 3)	$p.o.$	11	350	1	0	1.9209E-09	25.9563
	(6, 6)	$p. \ KKT$	16	359	0.25	0	1.6516E-09	57.4519
	(20, 20)	$p. \ KKT.$	5	105	0.0556	8.4143E-05	4.1327E-09	151.3305
PROBLEMAS PERTURBADOS USANDO RIP-FUNÇÃO $F(x)$								
Prob.	$(p, n)$	$(\alpha^*, \mathbf{x}^*)$	$it. \ ext.$	$it. \ int.$	$F$	$\ d^{k,tan}\ $	$\ C(s^*)\ $	Tempo
20	(3, 2)	$p.o.$	2	13	9.0003E-08	3.1101E-16	8.9382E-09	1.3969
		$p.KKT$	1	10	3.4483	2.2263E-05	2.7678E-07	0.6636
		$p.KKT$	2	16	3.4483	1.1102E-16	2.1468E-08	1.4589
	(6, 2)	$p.o.$	2	13	2.0667E-15	0	1.7514E-07	1.7688
	(100, 2)	$p.o.$	2	10	0	0	7.7715E-08	14.7777
	(300, 2)	$p.o.$	2	9	3.2883E-16	0	4.6134E-08	58.3963
21	(6, 6)	$p.o.$	1	3	0	0	9.9341E-09	0.5724
	(10, 10)	$p. \ max.$	1	7	1	1.3878E-17	5.1011E-09	2.5197
		$p.o.$	1	3	0	0	1.7881E-08	0.7427
	(20, 20)	$p.o.$	1	3	0	0	5.9605E-09	1.9950
22	(2, 2)	$p.o.$	1	8	1	0	1.5636E-08	0.5181
	(2, 2)	$p.o.$	3	71	1	0	7.4506E-09	5.9141
	(3, 3)	$p.KKT$	1	9	2	0	4.9671E-09	0.5581
	(6, 6)	$p. \ KKT$	1	28	5	0	1.9868E-08	2.7982
	(20, 20)	$p. \ KKT.$	1	35	5	1.3878E-17	4.5657E-08	29.0382

Tabela 6.12: Família de problemas com restrições - Método 2

Nas Tabelas 6.10, 6.11 e 6.12 observamos que os Métodos 1 e 2, junto com as suas variantes para o caso do PM ter restrições, são melhores que o Método 3, pois em todos os problemas testados chegam a um ponto KKT ou à solução do problema. Já a reformulação

KKT, além de obter pontos KKT ou chegar à solução do problema, em vários dos testes chega a pontos não viáveis, pontos não eficientes ou pontos que não correspondem a pontos KKT do problema reformulado.

Por outro lado vamos destacar que a variante RIP sobre o método RI tanto no Método 1 como no Método 2, teve melhor desempenho e, o M1-RIP consegue os melhores resultados no problema 20, em todos os testes feitos este método chega à solução do problema.

### 6.3.2 Experimentos sistematizados

Para os experimentos sistematizados rodamos os problemas 17, 18, e 19 com os três métodos e, os problemas 23, 24 e 25 usando os métodos 1 e 2 com as suas variantes e o método 3. A seguir são expostos os resultados obtidos para estas famílias.

Para estas rodadas escolhemos três conjuntos de valores diferentes para as dimensões dos vetores  $\alpha$  e  $x$ , ou seja para  $p$  e  $n$ . As escolhas foram as seguintes:

(1)  $p = 5$  e  $n = 10$ ,

(2)  $p = 10$  e  $n = 20$ ,

(3)  $p = 15$  e  $n = 30$ .

Para cada uma, foram gerados 100 pontos iniciais de três maneiras diferentes:

- (a)  $x_i^0 = \frac{1}{n} \quad \forall i = 1, \dots, n$ , e  $\alpha_i^0 \in (0, 1)$  aleatório  $\forall i = 1, \dots, p$  tal que  $\sum_{i=1}^p \alpha_i = 1$ ,
- (b)  $x_i^0 = \frac{i}{n} \quad \forall i = 1, \dots, n$ , e  $\alpha_i^0 \in (0, 1)$  aleatório  $\forall i = 1, \dots, p$  tal que  $\sum_{i=1}^p \alpha_i = 1$ ,
- (c)  $\alpha_i^0 = \frac{1}{n} \quad \forall i = 1, \dots, n$ , e  $x_i^0 \in (-10, 10)$  aleatório  $\forall i = 1, \dots, n$ .

Considerando a solução dos problemas 17 a 19 e 23 a 25, observamos que destas três formas diferentes de gerar pontos iniciais, os pontos do tipo (b) e (c) produzem instâncias com maior dificuldade que as gerados da forma (a). Isto é percebido nos resultados obtidos dos experimentos sistematizados. A seguir apresentamos os resultados obtidos nestes testes e mostramos a percentagem de pontos que satisfizeram a tolerância exigida para a otimalidade e a viabilidade, a percentagem de pontos que correspondem a pontos AGP e KKT. Também apresentamos a percentagem de pontos que satisfazem a tolerância da otimalidade mas que a inviabilidade pertence ao intervalo  $(10^{-04}, 10^{-02})$ . A

notação usada é a seguinte: **opt.** corresponde à percentagem de pontos que satisfazem a tolerância exigida para a otimalidade, **viab.** representa a percentagem de pontos que satisfazem a tolerância exigida para a viabilidade, **AGP** corresponde à percentagem de pontos que são pontos AGP, enquanto  $\sim\mathbf{AGP}$  corresponde à percentagem de pontos tais que  $\|C(x^k)\| \in (10^{-04}, 10^{-02})$  e satisfazem a tolerância exigida para otimalidade e, **KKT** representa a percentagem de pontos que satisfazem as condições KKT do subproblema da fase de otimização segundo os métodos de RI ou as condições KKT do subproblema a ser resolvido segundo a reformulação KKT.

As primeiras tabelas correspondem aos resultados dos problemas sem restrição e, em seguida, mostramos os resultados dos problemas com restrições.

DTLZ1												
(a)	$p = 5, n = 10,$				$p = 10, n = 20,$				$p = 15, n = 30,$			
	opt.	viab.	AGP	~AGP	opt.	viab.	AGP	~AGP	opt.	viab.	AGP	~AGP
M.1	98%	40%	38%	41%	100%	32%	32%	58%	98%	62%	60%	38%
M.2	100%	41%	41%	39%	100%	33%	33%	58%	95%	65%	63%	58%
		viab.	KKT			viab.	KKT			viab.	KKT	
M.3	100%	68%				99%	60%			92%	36%	
(b)	$p = 5, n = 10,$				$p = 10, n = 20,$				$p = 15, n = 30,$			
	opt.	viab.	AGP	~AGP	opt.	viab.	AGP	~AGP	opt.	viab.	AGP	~AGP
M.1	97%	24%	22%	55%	100%	37%	37%	56%	100%	28%	28%	68%
M. 2	100%	22%	22%	57%	99%	38%	37%	56%	100%	30%	30%	69%
		viab.	KKT			viab.	KKT			viab.	KKT	
M. 3	97%	53%				92%	41%			44%	16%	
(c)	$p = 5, n = 10,$				$p = 10, n = 20,$				$p = 15, n = 30,$			
	opt.	viab.	AGP	~AGP	opt.	viab.	AGP	~AGP	opt.	viab.	AGP	~AGP
M.1	99%	29%	29%	67%	100%	23%	22%	76%	100%	27%	27%	73%
M.2	99%	28%	28%	68%	100%	25%	25%	26%	100%	26%	26%	74%
		viab.	KKT			viab.	KKT			viab.	KKT	
M.3	97%	95%				90%	72%			92%	72%	
DTLZ2												
(a)	$p = 5, n = 10,$				$p = 10, n = 20,$				$p = 15, n = 30,$			
	opt.	viab.	AGP	~AGP	opt.	viab.	AGP	~AGP	opt.	viab.	AGP	~AGP
M.1	100%	100%	100%	0%	100%	100%	100%	0%	100%	100%	100%	0%
M.2	100%	100%	100%	0%	100%	100%	100%	0%	100%	100%	100%	0%
		viab.	KKT			viab.	KKT			viab.	KKT	
M.3	100%	50%				71%	4%			64%	5%	
(b)	$p = 5, n = 10,$				$p = 10, n = 20,$				$p = 15, n = 30,$			
	opt.	viab.	AGP	~AGP	opt.	viab.	AGP	~AGP	opt.	viab.	AGP	~AGP
M.1	98%	100%	98%	0%	100%	100%	100%	0%	99%	100%	99%	0%
M.2	98%	100%	98%	0%	99%	100%	99%	0%	99%	99%	98%	1%
		viab.	KKT			viab.	KKT			viab.	KKT	
M.3	19%	3%				18%	4%			17%	0%	
(c)	$p = 5, n = 10,$				$p = 10, n = 20,$				$p = 15, n = 30,$			
	opt.	viab.	AGP	~AGP	opt.	viab.	AGP	~AGP	opt.	viab.	AGP	~AGP
M.1	100%	100%	100%	0%	99%	100%	99%	0%	98%	100%	98%	0%
M.2	100%	100%	100%	0%	99%	100%	99%	0%	98%	100%	98%	0%
		viab.	KKT			viab.	KKT			viab.	KKT	
M.3	65%	22%				58%	14%			44%	11%	

**Tabela 6.13:** Tabela comparativa dos problemas DTLZ1 e DTLZ2 sem restrições (Problemas 17 e 18).



DTLZ3												
(a)	$p = 5, n = 10,$				$p = 10, n = 20,$				$p = 15, n = 30,$			
	opt.	viab.	AGP	~AGP	opt.	viab.	AGP	~AGP	opt.	viab.	AGP	~AGP
M.1	100%	20%	20%	74%	100%	26%	26%	60%	100%	53%	53%	19%
M.2	100%	21%	21%	74%	100%	26%	26%	61%	100%	53%	53%	18%
		viab.	KKT			viab.	KKT			viab.	KKT	
M.3		93%	17%			80%	6%			37%	1%	

(b)	$p = 5, n = 10,$				$p = 10, n = 20,$				$p = 15, n = 30,$			
	opt.	viab.	AGP	~AGP	opt.	viab.	AGP	~AGP	opt.	viab.	AGP	~AGP
M.1	99%	27%	26%	57%	98%	47%	45%	36%	100%	50%	50%	39%
M.2	98%	30%	30%	53%	100%	45%	45%	39%	100%	50%	50%	45%
		viab.	KKT			viab.	KKT			viab.	KKT	
M.3		94%	23%			66%	4%			55%	2%	

(c)	$p = 5, n = 10,$				$p = 10, n = 20,$				$p = 15, n = 30,$			
	opt.	viab.	AGP	~AGP	opt.	viab.	AGP	~AGP	opt.	viab.	AGP	~AGP
M.1	100%	0%	0%	45%	100%	4%	4%	59%	100%	0%	0%	82%
M.2	100%	0%	0%	45%	100%	3%	3%	59%	100%	1%	1%	82%
		viab.	KKT			viab.	KKT			viab.	KKT	
M.3		76%	13%			61%	15%			50%	7%	

Tabela 6.14: Tabela comparativa do problema DTLZ3 sem restrições (problema 19).

DTLZ1 - com Restrição												
(a)	$p = 5, n = 10,$				$p = 10, n = 20,$				$p = 15, n = 30,$			
	opt.	viab.	AGP	~AGP	opt.	viab.	AGP	~AGP	opt.	viab.	AGP	~AGP
M.1-R.I.	100%	34%	34%	13%	100%	80%	80%	2%	100%	65%	65%	35%
M.1-R.I.P.	96%	76%	74%	8%	99%	97%	96%	3%	100%	93%	93%	1%
M.2-R.I.	100%	36%	36%	27%	100%	80%	80%	2%	100%	63%	63%	23%
M.2-R.I.P.	95%	79%	76%	4%	98%	92%	90%	0%	97%	91%	88%	2%
		viab.	KKT			viab.	KKT			viab.	KKT	
M.3	100%	43%			95%	25%			93%	21%		
(b)	$p = 5, n = 10,$				$p = 10, n = 20,$				$p = 15, n = 30,$			
	opt.	viab.	AGP	~AGP	opt.	viab.	AGP	~AGP	opt.	viab.	AGP	~AGP
M.1-R.I.	100%	30%	30%	17%	100%	45%	45%	1%	100%	17%	17%	6%
M.1-R.I.P.	99%	92%	92%	0%	100%	94%	94%	6%	99%	94%	93%	0%
M.2-R.I.	100%	29%	29%	18%	100%	46%	46%	1%	100%	17%	17%	6%
M.2-R.I.P.	100%	92%	92%	0%	96%	93%	90%	6%	98%	94%	92%	1%
		viab.	KKT			viab.	KKT			viab.	KKT	
M.3	0%	0%			0%	0%			0%	0%		
(c)	$p = 5, n = 10,$				$p = 10, n = 20,$				$p = 15, n = 30,$			
	opt.	viab.	AGP	~AGP	opt.	viab.	AGP	~AGP	opt.	viab.	AGP	~AGP
M.1-R.I.	99%	19%	19%	16%	100%	21%	21%	25%	100%	28%	28%	35%
M.1-R.I.P.	100%	71%	71%	15%	100%	77%	77%	23%	100%	68%	68%	19%
M.2-R.I.	100%	20%	20%	16%	100%	21%	21%	25%	100%	27%	27%	35%
M.2-R.I.P.	100%	71%	71%	16%	100%	77%	77%	23%	98%	67%	65%	19%
		viab.	KKT			viab.	KKT			viab.	KKT	
M.3	2%	1%			1%	0%			0%	0%		

Tabela 6.15: Tabela comparativa do problema DTLZ1 com restrições (Problema 23).

DTLZ2 - com Restrição												
(a)	$p = 5, n = 10,$				$p = 10, n = 20,$				$p = 15, n = 30,$			
	opt.	viab.	AGP	~AGP	opt.	viab.	AGP	~AGP	opt.	viab.	AGP	~AGP
M.1-R.I.	100%	5%	5%	95%	100%	0%	0%	100%	100%	2%	2%	98%
M.1-R.I.P.	100%	100%	100%	0%	100%	98%	98%	2%	100%	100%	100%	0%
M.2-R.I.	100%	5%	5%	95%	100%	0%	0%	100%	100%	2%	2%	98%
M.2-R.I.P.	100%	100%	100%	0%	100%	99%	99%	1%	100%	100%	100%	0%
		viab.	KKT			viab.	KKT			viab.	KKT	
M.3		92%	71%			5%	2%			3%	0%	
(b)	$p = 5, n = 10,$				$p = 10, n = 20,$				$p = 15, n = 30,$			
	opt.	viab.	AGP	~AGP	opt.	viab.	AGP	~AGP	opt.	viab.	AGP	~AGP
M.1-R.I.	100%	5%	5%	95%	100%	0%	0%	100%	100%	0%	0%	100%
M.1-R.I.P.	100%	100%	100%	0%	100%	97%	97%	3%	100%	99%	99%	1%
M.2-R.I.	100%	5%	5%	95%	100%	1%	1%	99%	100%	0%	0%	100%
M.2-R.I.P.	100%	100%	100%	0%	100%	96%	96%	4%	100%	99%	99%	1%
		viab.	KKT			viab.	KKT			viab.	KKT	
M.3		78%	11%			22%	10%			0%	0%	
(c)	$p = 5, n = 10,$				$p = 10, n = 20,$				$p = 15, n = 30,$			
	opt.	viab.	AGP	~AGP	opt.	viab.	AGP	~AGP	opt.	viab.	AGP	~AGP
M.1-R.I.	100%	88%	88%	12%	100%	91%	91%	9%	100%	93%	93%	7%
M.1-R.I.P.	100%	98%	98%	2%	100%	39%	39%	61%	100%	100%	100%	0%
M.2-R.I.	100%	88%	88%	12%	100%	91%	91%	9%	100%	93%	93%	7%
M.2-R.I.P.	100%	98%	98%	2%	100%	40%	40%	60%	100%	100%	100%	0%
		viab.	KKT			viab.	KKT			viab.	KKT	
M.3		10%	2%			0%	0%			0%	0%	

Tabela 6.16: Tabela comparativa do problema DTLZ2 com restrições (Problema 24).

DTLZ3 - com Restrição												
(a)	$p = 5, n = 10,$				$p = 10, n = 20,$				$p = 15, n = 30,$			
	opt.	viab.	AGP	$\sim$ AGP	opt.	viab.	AGP	$\sim$ AGP	opt.	viab.	AGP	$\sim$ AGP
M.1-R.I.	100%	37%	37%	1%	100%	84%	84%	1%	100%	29%	29%	2%
M.1-R.I.P.	100%	59%	59%	11%	100%	83%	83%	1%	100%	93%	93%	0%
M.2-R.I.	100%	36%	36%	1%	100%	85%	85%	1%	100%	29%	29%	2%
M.2-R.I.P.	100%	60%	60%	11%	97%	83%	80%	1%	90%	93%	84%	0%
		viab.	KKT			viab.	KKT			viab.	KKT	
M.3		92%	7%			88%	3%			13%	0%	
(b)	$p = 5, n = 10,$				$p = 10, n = 20,$				$p = 15, n = 30,$			
	opt.	viab.	AGP	$\sim$ AGP	opt.	viab.	AGP	$\sim$ AGP	opt.	viab.	AGP	$\sim$ AGP
M.1-R.I.	96%	41%	41%	1%	100%	9%	9%	3%	99%	9%	9%	0%
M.1-R.I.P.	100%	95%	95%	5%	100%	99%	99%	1%	100%	100%	100%	0%
M.2-R.I.	99%	41%	41%	6%	99%	9%	9%	3%	98%	11%	11%	0%
M.2-R.I.P.	99%	95%	94%	5%	100%	99%	99%	1%	97%	100%	97%	0%
		viab.	KKT			viab.	KKT			viab.	KKT	
M.3		0%	0%			15%	0%			8%	0%	
(c)	$p = 5, n = 10,$				$p = 10, n = 20,$				$p = 15, n = 30,$			
	opt.	viab.	AGP	$\sim$ AGP	opt.	viab.	AGP	$\sim$ AGP	opt.	viab.	AGP	$\sim$ AGP
M.1-R.I.	100%	74%	74%	0%	100%	70%	70%	0%	99%	77%	77%	1%
M.1-R.I.P.	100%	88%	88%	12%	100%	90%	90%	10%	100%	91%	91%	9%
M.2-R.I.	100%	74%	74%	0%	100%	70%	70%	0%	100%	78%	78%	3%
M.2-R.I.P.	100%	88%	88%	12%	100%	90%	90%	10%	100%	91%	91%	9%
		viab.	KKT			viab.	KKT			viab.	KKT	
M.3		0%	0%			3%	0%			0%	0%	

**Tabela 6.17:** Tabela comparativa do problema DTLZ3 com restrições (Problema 25).

Das Tabelas 6.13–6.17 podemos observar que os Métodos 1 e 2, junto com as suas variantes para o caso de ter restrições o **P.M.**, a otimalidade é atingida em quase todos os testes, entanto a viabilidade não, mas obtemos pontos *aproximadamente AGP*, isto é  $\sim$ **AGP**. Estes últimos pontos satisfazem a tolerância da otimalidade mas a inviabilidade pertence ao intervalo  $(10^{-04}, 10^{-02})$ , isto é porque os métodos de RI param por cumprir o outro critério de convergência, aquele no qual a distancia entre o ponto intermediário e o obtido na iteração anterior é menor que uma certa tolerância.

### Tabelas comparativas da natureza dos pontos obtidos nos experimentos sistematizados

A seguir apresentamos quatro tabelas que contêm em percentagem, a natureza dos pontos obtidos com cada método utilizado para resolver os problemas DTLZ1, DTLZ2 e DTLZ3 com e sem restrições. Esta informação é apresentada de duas maneiras. Nas primeiras duas tabelas separamos os pontos por dimensão ( $p=5$ ,  $n=10$ ;  $p=10$ ,  $n=20$  e  $p=15$ ,  $n=30$ ) e as outras duas contêm a informação separada pelo tipo de pontos iniciais gerados ((a), (b) e (c)).

	Método 1	Método 2	Método 3
$(p, n)$	PONTOS AGP		PONTOS KKT
(5, 10)	48.11%	48.88%	38.22%
(10, 15)	51.66%	51.88%	24.44%
(15, 30)	57.22%	57.66%	16.66%
<b>Total</b>	52.33%	52.81%	26.44%

**Tabela 6.18:** Tabela comparativa separando por dimensão. Problemas sem restrições (17, 18 e 19).

	Método 1	Método 2	Método 3
Inicialização	PONTOS AGP		PONTOS KKT
(a)	58.77%	59.88%	27.44%
(b)	56.11%	56.55%	16.22%
(c)	42.11%	42.22%	35.66%

**Tabela 6.19:** Tabela comparativa agrupando as dimensões e separando pela inicialização adotada. Problemas sem restrições (17, 18 e 19).

	Método 1		Método 2		Método 3
	RI	RIP	RI	RIP	
$(p, n)$	PONTOS AGP				PONTOS KKT
(5, 10)	37%	86.33%	37.11%	86.55%	15%
(10, 15)	44.44%	85.88%	44.77%	84.55%	4.44%
(15, 30)	35.55%	93%	35.55%	90.66%	2.33%
<b>Total</b>	39%	88.40%	39.14%	87.25%	7.25%

**Tabela 6.20:** Tabela comparativa separando por dimensão. Problemas com restrições (23, 24 e 25).

	Método 1		Método 2		Método 3
	RI	RIP	RI	RIP	
Inicialização	PONTOS AGP				PONTOS KKT
(a)	37.33%	88.44%	37.33%	86.33%	19.11%
(b)	17.33%	96.55%	17.66%	95.44%	2.33%
(c)	62.33%	80.22%	62.44%	80%	0.33%

**Tabela 6.21:** Tabela comparativa agrupando as dimensões e separando pela inicialização adotada. Problemas com restrições (23, 24 e 25).

Das Tabelas 6.18–6.21 podemos observar que com os Métodos 1 e 2 obtemos melhores resultados que com o Método 3. No caso dos problemas sem restrições, vemos da Tabela 6.18, que quando a dimensão dos problemas aumenta, a quantidade de pontos AGP é maior, enquanto no Método 3, a quantidade de pontos KKT diminui. Se olharmos para a Tabela 6.19, também é possível perceber que os Métodos 1 e 2 são melhores que o Método 3, para os três tipos de pontos gerados.

No caso dos problemas com restrições observamos, como no caso anterior, que com os Métodos 1 e 2 obtemos melhores resultados do que com a reformulação KKT. Para estes problemas temos também as variantes dos Métodos 1 e 2: M1-RIP e M2-RIP. Das tabelas vemos que os resultados usando tais variantes são muito melhores.

Considerando os experimentos sistematizados e as suas soluções, pode-se pensar que a restrição não faz diferença no problema, pois para o caso de estar ou não presente,

a solução é sempre a origem, e assim a restrição fica inativa. Mas a restrição faz uma grande diferença nos resultados obtidos, podemos ver na Tabela 6.20 que quando aumenta a dimensão, com os método que usam RI os resultados são melhores e com grande superioridade com as suas variantes RIP, no entanto com a reformulação KKT os resultados pioram quando aumenta a dimensão. Da tabela 6.21 podemos dizer novamente que os resultados usando RIP são muito melhores em relação aos outros métodos.

Outro ponto a destacar dos problemas DTLZ com restrições é a convergência para seus extremos absolutos. Tais pontos foram identificados dentre os pontos AGP obtidos, com valor de função objetivo menor que 0.001 para minimizadores globais, ou entre 0.999 e 1.001 para maximizadores globais. O bom desempenho dos métodos que usam RIP é evidenciado nas tabelas apresentadas a seguir.

$(p, n)$	Método 1		Método 2		Método 3
	RI	RIP	RI	RIP	
MINIMO					
(5, 10)	0.11%	30%	0.11%	29.88%	0.11%
(10, 15)	0.11%	31.33%	0.11%	31.11%	—
(15, 30)	3.33%	37.77%	3.22%	37%	0.33%
MÁXIMO					
(5, 10)	31.44%	11.55%	31.11%	11.55%	—
(10, 15)	28.44%	18.66%	28.44%	17.66%	—
(15, 30)	25.55%	21.66%	25.77%	20.55%	—

**Tabela 6.22:** Resultados por dimensão. Problemas com restrições (23, 24 e 25)

	Método 1		Método 2		Método 3
	RI	RIP	RI	RIP	
Inicialização	MÍNIMO				
(a)	3.22%	9.77%	3.22%	9.33%	0.33%
(b)	0.11%	49.66%	0.11%	49%	—
(c)	0.22%	39.66%	0.22%	39.66%	0.11%
	MÁXIMO				
(a)	15.77%	19.22%	15.55%	17.88%	—
(b)	14.88%	11.55%	15%	11%	—
(c)	54.77%	21.11%	54.77%	20.88%	—

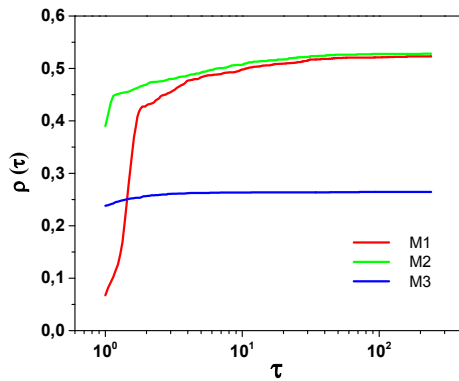
**Tabela 6.23:** Resultados por ponto gerado. Problemas com restrições (Problemas 23, 24 e 25)

Novamente podemos observar a superioridade dos métodos que usam RIP sobre os outros métodos. Das Tabelas 6.22 e 6.23 vemos que com esses métodos foi possível atingir a extremos absolutos dos problemas, e em maior percentagem ao mínimo absoluto. Mais ainda vamos destacar que na tabela de pontos gerados, os casos (b) e (c), os quais mencionamos produzir instâncias mais difíceis, obtivemos os melhores resultados com RIP, enquanto o desempenho do Método 3 foi o pior.

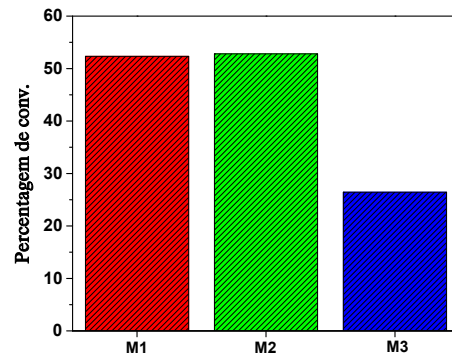
### 6.3.3 Gráficos de desempenho dos experimentos sistematizados

A seguir comparamos o desempenho dos métodos implementados na resolução dos testes sistematizados (problemas DTLZ) mediante vários gráficos. Estes são agrupados segundo o problema multiobjetivo seja restrito ou não. Mostramos o perfil de desempenho dos tempos do CPU utilizado pelos três métodos para satisfazer os nossos critérios de convergência, conforme apresentado por Dolan e Moré [13]. Aqui, a curva desenhada representa a função de distribuição acumulativa de desempenho do tempo do CPU. Também apresentamos as percentagens de convergência de cada método.

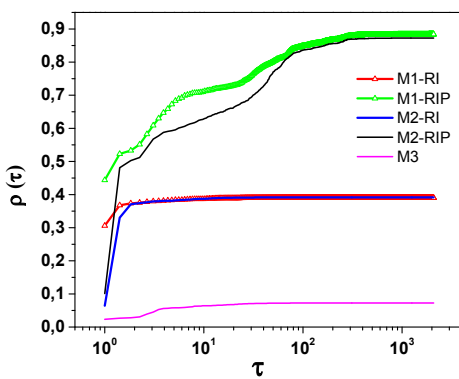




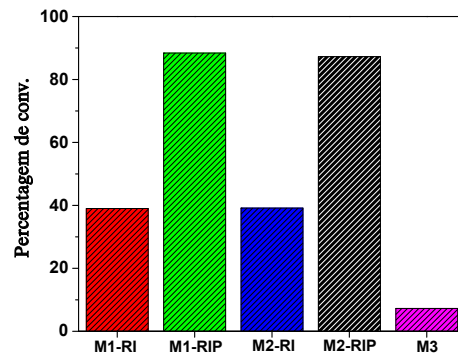
(a) Perfil de desempenho do tempo



(b) Percentagem de convergência

**Figura 6.1:** Gráficos para problemas DTLZ sem restrições (Problemas 17, 18 e 19).

(a) Perfil de desempenho do tempo



(b) Percentagem de convergência

**Figura 6.2:** Gráficos para problemas DTLZ com restrições (Problemas 23, 24 e 25)

Da Figura 6.1 observamos que se estamos interessados no método que possa resolver 40% dos problemas com a maior eficiência, então Método 2 se destaca, enquanto o Método 1 é o menos vantajoso. Porém o desempenho do Método 2 torna-se muito mais competitivo se estendermos  $\tau$  de interesse para 2.5. Se estamos interessados na probabilidade para resolver os problemas exitosamente, então os Métodos 1 e 2 são os melhores. Da Figura 6.2 podemos ver que o M1-RIP corresponde ao método mais eficiente dos métodos testados e o Método 3, corresponde ao método com menor eficiência. Por outro lado, enquanto o valor de  $\tau$  aumenta o M2-RIP começa a se destacar, e em termos da robustez, este último junto com o M1-RIP são os melhores. Destes gráficos podemos dizer que no caso dos problemas apresentarem restrições no problema multiobjetivo, o melhor método para

estes problemas foi M1-RIP, enquanto no caso irrestrito o melhor foi o Método 2.

#### 6.3.4 Comentários sobre os experimentos

Para comparar o desempenho dos três métodos propostos, foram rodados experimentos preliminares e experimentos sistematizados. Dos testes preliminares foi possível observar que com os Métodos 1 e 2 obtivemos melhores resultados do que com o Método 3. Nos testes sistematizados, confirmamos estes resultados por meio dos gráficos de perfil de desempenho e da percentagem de convergência junto com as tabelas comparativas dos pontos obtidos. Para os problemas que possuem restrições além da caixa, pudemos ver que além de obter melhores resultados com os Métodos 1 e 2 com respeito à reformulação KKT, as suas variantes M1-RIP e M2-RIP fazem uma grande diferença na solução em relação ao método original. Além de aumentar a percentagem de pontos AGP obtidos, é possível verificar que com essas variantes achamos o máximo e o mínimo absoluto para alguns dos problemas. Já com a reformulação KKT, em vários dos testes nem foi possível achar pontos viáveis.

---

---

## Capítulo 7

### Conclusões

O tipo de problema apresentado nesta tese corresponde a uma classe importante de problemas de tomada de decisões. Quando as funções objetivo e as restrições do **(PM)** são convexas, é possível transformar o problema multiobjetivo em um problema escalar adequado com pesos. Escolhendo os pesos maiores que algum  $\epsilon$  estamos permitindo que todos os objetivos influenciem na tomada de decisões, isto é, nenhum dos objetivos é negligenciado. Com o problema transformado e, considerando a convexidade nas restrições, temos ferramentas teóricas para achar os pontos propriamente eficientes do problema de segundo nível e assim, achar o conjunto factível do problema de primeiro nível. Também sob certas hipóteses foi possível provar que este conjunto é convexo, o que permite garantir a convergência dos métodos propostos.

Dos experimentos preliminares, foi possível verificar que no caso em que o **(PM)** não apresente restrições, e as funções objetivo sejam convexas, os três métodos convergem, convergindo em menor tempo o Método 3. No entanto, no caso do problema multiobjetivo apresentar restrições convexas, os métodos *R.I.* convergem ao ótimo e, em dois casos atingem um ponto de máximo local. Além disso, dos gráficos de perfil de desempenho, vemos que os métodos 1 e 2 têm desempenhos semelhantes, ganhando em tempo o Método 2. Já o método que usa as condições KKT obtém pontos que não correspondem a extremos do problema.

Para mostrar a robustez dos métodos, testamos problemas que não cumprem as hipóteses dos Teoremas 2.1 e 2.2, não cumprem a diferenciabilidade das funções objetivo do **(PM)** (problema 9, com  $n = 2$ ), e que o conjunto eficiente não corresponde a uma curva diferenciável (problema 13). Os Métodos 1 e 2 atingem a solução nestes experimentos, no entanto o Método 3 não converge para todas as inicializações e, no problema 9

não consegue pontos factíveis.

Nos experimentos sistematizados, as funções coordenadas dos problemas multiobjetivo propostos não correspondem a funções unimodais, de fato elas tem muitos mínimos e máximos locais. Assim, com os Métodos 1 e 2, na fase de restauração obtivemos soluções locais do problema de programação vetorial e, portanto podemos dizer que os pontos obtidos correspondem a pontos eficientes locais, isto é,  $\mathbf{x}$  é um ponto eficiente local se  $\exists \delta > 0$  tal que  $\mathbf{x}$  é eficiente em  $\mathcal{B}(\mathbf{x}, \delta)$ . Logo, os pontos obtidos na fase de otimização correspondem a máximos locais, mínimos locais ou simplesmente a pontos AGP do problema de primeiro nível do problema reformulado. Já com o Método 3 os pontos obtidos correspondem a pontos KKT do problema proposto a resolver, os quais não temos garantia de ser soluções do problema multiobjetivo.

No caso do problema multiobjetivo apresentar restrições, foi possível verificar com um exemplo simples (problema (1.3)) e o problema 20 dos experimentos testados, que no caso de utilizar os Métodos 1 e 2, é possível ficar preso num ponto indesejável quando se usa *Restauração Exata* na fase de restauração. A estratégia proposta para contornar essa dificuldade foi utilizar RI nessa fase e então perturbar a restrição de complementaridade associada ao problema de nível inferior. Dos experimentos testados, verificamos que usando esta estratégia nos Método 1 e 2 obtivemos os melhores resultados. Isto foi testado primeiro com o problema 20 e também nos experimentos sistematizados. Mais ainda, com os testes sistematizados observamos que com o Método 3, na maioria dos casos, não foi possível achar pontos factíveis para os problemas quando dificultamos os pontos iniciais e aumentamos a dimensão do problema. Podemos dizer que nestes experimentos, as duas propostas com RI, no caso de problemas irrestritos, e as duas propostas de RIP para o caso de problema com restrições, obtiveram o melhor desempenho na resolução de problemas em dois níveis-multiobjetivo.

Concluimos, portanto, que utilizar RI é mais vantajoso que utilizar a reformulação KKT, e no caso de haver restrições é preferível utilizar RIP, pois é possível contornar os pontos indesejáveis e chegar à solução do problema. Também a estratégia de transformar o (PM) em um problema escalar adequado é muito favorável, pois permite abordar o problema vetorial com muitos objetivos, o que em geral não é possível pela complexidade que isto envolve e também permite utilizar métodos de programação não linear para resolver os problemas vetoriais diferentemente da maioria das propostas que utilizam

modelos heurísticos para resolvê-los.

Como trabalho futuro, gostaríamos de analisar o desempenho destes métodos quando o nível superior é também um problema multiobjetivo. Também pensamos em investigar as propriedades da homotopia usada na perturbação de forma exaustiva em outros problemas.

# Referências Bibliográficas

- [1] ALLENDE G. B., STILL G., “Solving bilevel programs with KKT-approach”, *Mathematical Programming*, Ser. A 138 (2013), 309-332.
- [2] ANDREANI R., CASTRO S. L. DE C., CHELA J. L., FRIEDLANDER A., SANTOS S. A., “An inexact-restoration method for nonlinear bilevel programming problems”, *Computational Optimization and Applications*, 43 (2009), 307-328.
- [3] BENSON B., “An improved definition of proper efficiency for vector maximization with respect to cones”, *J. Math. Anal. Appl.*, 71 (1979), 232-241.
- [4] BIRGIN E. G., MARTÍNEZ J. M., RAYDAN M., “Nonmonotone spectral projected gradients methdos on convexs sets”, *SIAM J. Optim.*, 10 (2000), 1196-1211.
- [5] BIRGIN E. G., MARTÍNEZ J. M., RAYDAN M., “SPG - Software for convex-constrained optimization”, *ACM Trans. Math. Software*, 27 (2001), 340-349.
- [6] BIRGIN E. G., MARTÍNEZ J. M., RAYDAN M., “Inexact Spectral Projected Gradient methods on convexs sets”, *IMA J. Numer. Anal.*, 23 (2003), 539-559.
- [7] BORWEIN J., “Proper efficient points for maximization with respect to cones”, *SIAM J. Control Optimization*, 15 (1977), 57-63.
- [8] BURACHIK S. R., RIZVI S. S., “On Weak and Strong Kuhn-Tucker Conditions for Smooth Multiobjective optimization”, *Journal of Optimization Theory and Applications*, 155(3)(2012), 923-940.
- [9] CHANKONG V., HAIMES Y. Y., *Multiobjective Decision Making: Theory and Methodology*, North Holland Series in System Science and Engineering; volume 8, 1983.

- [10] DEB K., THIELE L., LAUMANN S., ZITZLER E., “Scalable Test Problems for Evolutionary Multi-Objective Optimization”, *Kanpur Genetic Algorithms Lab., Indian Inst. Technol., Report 2* (2001), 001 001.
- [11] DEMPE S., *Foundations of Bilevel Programming*, Kluwer Academic Publishers, New York, Boston, Dordrecht, London, Moscow, 2002.
- [12] DEMPE S., “Annotated bibliography on bilevel programming and mathematical problems with equilibrium constraints”, *Optimization* 52 (2003), 333-359.
- [13] DOLAN E. D., MORÉ J. J., “Benchmarking Optimization, Software with Performance Profiles”, *Mathematical Programming* 91.2 (2002), 201-213.
- [14] EDGEWORTH F. Y., *Mathematical Psychics*, UNIVERSITY MICROFILMS INTERNATIONAL (OUT-OF-PRINT BOOKS ON DEMAND), 1987 (THE ORIGINAL EDITION IN 1881).
- [15] EHRGOTT M., *Multicriteria Optimization*, 2nd ed, Springer, Berlin, Heidelberg, 2000.
- [16] EICHEFELDER G., *Adaptive Scalarization Methods in Multiobjective Optimization*, Springer-Verlag Berlin, Heidelberg, 2008.
- [17] FLETCHER R., *Practical Methods of Optimization*, 2nd ed, Willey, Chichester, UK, 1987.
- [18] FISCHER A., FRIEDLANDER A., “A new line search inexact restoration approach for nonlinear programming”, *Comp. Optim. Appl.* 46 (2010), 333-346.
- [19] GEOFFRION A. M., “Proper efficiency and theory of vector maximization”, *Operations Research* 15 (1968), 39-54.
- [20] GOMES-RUGGIERO M. A., MARTÍNEZ J. M., SANTOS S. A., “Spectral projected gradient method with inexact restoration for minimization with nonconvex constraints”, *SIAM. Sci. Comput.* 31 (2009), 1628-1652.
- [21] HUBAND S., BARONE L., WHILE L., HINGSTON P., “A scalable multiobjective test problem Toolkit”, *Computer Science-Berlin: Springer-Verlag*, Mar. 2005, vol. 3410 , *Proc. Evolutionary Multi-Criterion Optimization: 3rd Int. Conf.* (2005), 280-294.

- [22] HUBAND S., HINGSTON P., BARONE L., WHILE L., “A review of Multiobjective test Problems and Scalable test Problem Toolkit”, *IEEE Transactions on Evolutionary Computation*, 10(5) (2006), 477-506.
- [23] ISERMANN M., “Proper efficiency and linear vector maximum problems”, *Operations Research* 22 (1976), 189-191.
- [24] JAHN J., *Vector Optimization: Theory, Application, and Extensions*, 2nd ed, Springer Verlag Berlin, Heidelberg, 2011.
- [25] MIETTINEN KAISA M., *Nonlinear Multiobjective Optimization*, Kluwer, Boston, Dordrecht, 1999.
- [26] KUHN H. W., TUCKER A. W., *Nonlinear Programming*, in Proceedings of the Second Berkeley Symposium on Mathematics, Statistics, and Probability, University of California Press, Berkeley, (1951) pp. 481-492.
- [27] LUENBERGER G., *Linear and Nonlinear Programming*, Addison Wesley, Reading, MA, 1989.
- [28] MANGASARIAN O.L., *Nonlinear Programming*, SIAM, Philadelphia, 1994.
- [29] MARTÍNEZ J. M., SANTOS S. A., *Métodos Computacionais de Otimização*, Departamento de Matemática Aplicada, IMECC, 1995.
- [30] MARTÍNEZ J. M., SVAITER B. F., “A practical optimality condition without constraint qualifications for nonlinear programming”, *Journal of Optimization Theory and Applications* 118 (2003), pp. 117-133.
- [31] MARTÍNEZ J. M., PILOTA E., “Inexact-Restoration Algorithm for Constrained Optimization”, *Journal of Optimization Theory and Applications* 104 (2000), pp. 135-163.
- [32] MARTÍNEZ J. M., “Inexact-Restoration Method with Lagrangian Tangent Decrease and New Merit Function for Nonlinear Programming”. *Journal of Optimization Theory and Applications* 111 (2001), pp. 39-58.
- [33] NOCEDAL J., WRIGHT S. J., *Numerical Optimization*, Second Edition, Springer Series in Operations Research, Springer Verlag, 2006.



- 
- [34] PARETO V. *Cours d'Economie Politique* Libraire Droz, Genève, 1964 (the first edition in 1896).
- [35] ZITZLER E., DEB K., THIELE L., Comparison of Multiobjective Evolutionary Algorithms: Empirical Results. *Evolutionary Computation*, 8(2) (2000), pp. 173-195.